



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS ARARANGUÁ
CENTRO DE CIÊNCIAS, TECNOLOGIAS E SAÚDE
TECNOLOGIAS DA INFORMAÇÃO E COMUNICAÇÃO

Muriel Ricardo Ramos Andrigueto

**Método de Visualização de Dados para Geração de *Insights* na Gestão de
Ideias**

Araranguá

2023

Muriel Ricardo Ramos Andrigueto

**Método de Visualização de Dados para Geração de *Insights* na Gestão de
Ideias**

Trabalho de Conclusão de Curso submetido ao curso de Tecnologias da Informação e Comunicação do Centro de Ciências, Tecnologias e Saúde da Universidade Federal de Santa Catarina como requisito para a obtenção do título de Bacharel em Tecnologias da Informação e Comunicação.

Orientador(a): Prof.(a) Dr.(a) Marina Carradore Sérgio.

Araranguá

2023

ANDRIGUETO, Muriel Ricardo Ramos

Método de Visualização de Dados para Geração de Insights na
Gestão de Ideias / Muriel Ricardo Ramos ANDRIGUETO ;
orientadora, Marina Carradore Sérgio, 2023.
65 p.

Trabalho de Conclusão de Curso (graduação) - Universidade
Federal de Santa Catarina, Campus Araranguá, Graduação em
Tecnologias da Informação e Comunicação, Araranguá, 2023.

Inclui referências.

1. Tecnologias da Informação e Comunicação. 2. Visualização
de Dados. 3. Gestão de Ideias. 4. Inovação. I. Sérgio, Marina
Carradore. II. Universidade Federal de Santa Catarina. Graduação
em Tecnologias da Informação e Comunicação. III. Título.

Muriel Ricardo Ramos Andrigueto

Método de Visualização de Dados para Geração de *Insights* na Gestão de Ideias

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do título de “Bacharel em Tecnologias da Informação e Comunicação” e aprovado em sua forma final pelo Curso Tecnologias da Informação e Comunicação.

Araranguá, 05 de dezembro de 2023.

Prof. Fernando Jose Spanhol, Dr.
Coordenador do Curso

Banca examinadora

Prof.(a) Marina Carradore Sérgio, Dr.(a)
Orientador(a)

Prof. Fabrício Herpich, Dr.
Universidade Federal de Santa Catarina

Prof. Alexandre Leopoldo Gonçalves, Dr.
Universidade Federal de Santa Catarina

Araranguá, 2023

AGRADECIMENTOS

Minha primeira palavra de gratidão é para Deus, que me sustentou com Sua força e sabedoria ao longo desta jornada. A Ele, dedico toda a honra e glória por esta conquista.

A meus pais, Charlene e Alex, estendo minha profunda gratidão. O incentivo que me deram para seguir os estudos foram fundamentais para que eu chegasse até aqui. Esta conquista é um reflexo do amor e apoio que sempre me proporcionaram.

Meus sinceros agradecimentos à minha orientadora, Marina Carradore Sérgio, cuja orientação e apoio foram fundamentais para a realização deste trabalho. Sua disposição e comprometimento com a excelência foram vitais para a qualidade desse projeto. Sou grato pela sua paciência e por todo o conhecimento compartilhado.

Agradeço de coração ao pessoal do VIVOS, que se tornou uma verdadeira família para mim. Sou grato por cada oração e por me ajudarem a transformar este trabalho em mais um testemunho de conquista do nosso grupo de convívio.

A ela, Marina Barbosa, a "Ma", dedico uma gratidão especial. Obrigado por cada palavra de incentivo e por estar ao meu lado em todos os momentos. Sua capacidade de tornar até mesmo os momentos difíceis mais leves foram fundamentais para que eu chegasse até aqui.

Por fim, expresso minha gratidão a todos os professores e colegas da instituição que tive a oportunidade de conhecer e que, de alguma forma, contribuíram para esta graduação.

RESUMO

Na era digital, onde a informação cresce exponencialmente, a visualização de dados se torna uma técnica crucial para transformar informações textuais complexas em *insights* acionáveis, essencial para a análise de ideias inovadoras. Este trabalho investiga a aplicação de técnicas de visualização de dados em ideias coletadas do site "Cool Business Ideas", com o objetivo de revelar tendências, padrões e conexões ocultas, contribuindo para uma gestão de ideias mais eficaz e inovadora. Empregando uma abordagem quantitativa, o estudo analisou 904 ideias de negócios na categoria "Tecnologia" do período de 2004 a 2023, utilizando TF-IDF para vetorização e K-Means para agrupamento, complementados pela Análise de Componentes Principais (PCA) para redução de dimensionalidade. Os resultados apontam para um foco crescente em temas como inteligência artificial, sustentabilidade e soluções autônomas, destacando a importância de colaborações interdisciplinares e a integração entre áreas como IA e práticas sustentáveis. Conclui-se que a visualização de dados é uma ferramenta poderosa para a gestão e compreensão de ideias inovadoras, oferecendo *insights* valiosos para estratégias de negócios e inovação, e enfatizando a necessidade de abordagens que combinam tecnologia avançada com responsabilidade ambiental e social.

Palavras-chave: Visualização de Dados; Gestão de Ideias; Inovação.

ABSTRACT

In the digital era, characterized by the exponential growth of information, data visualization emerges as a crucial technique for transforming complex textual information into actionable insights, essential in the analysis of innovative ideas. This study investigates the application of data visualization techniques to ideas gathered from the "Cool Business Ideas" website, aiming to uncover trends, patterns, and hidden connections, and thereby contributing to more effective and innovative idea management. Employing a quantitative approach, the research analyzed 904 business ideas in the "Technology" category from 2004 to 2023, using TF-IDF for data vectorization and K-Means for clustering, supplemented by Principal Component Analysis (PCA) for dimensionality reduction. The results indicate an increasing focus on themes such as artificial intelligence, sustainability, and autonomous solutions, emphasizing the importance of interdisciplinary collaborations and the integration of areas like AI and sustainable practices. The study concludes that data visualization is a powerful tool for managing and understanding innovative ideas, offering valuable insights for business strategies and innovation, and underscoring the need for approaches that combine advanced technology with environmental and social responsibility.

Keywords: Data Visualization; Idea Management; Innovation.

LISTA DE FIGURAS

Figura 1 – Fluxograma de Processo de Gestão de Ideias: Coleta, Avaliação, Seleção e Implementação.....	21
Figura 2 – Página da categoria 'Tecnologia' no site CoolBusinessIdeas.com.....	33
Figura 3 – Extração de dados com a extensão 'Instant Data Scraper'.....	34
Figura 4 – Antes e depois da limpeza e preparação dos dados.....	38
Figura 5 – Visualização dos clusters após execução do K-Means na matriz TF-IDF	40
Figura 6 – Gráfico de dispersão de clusters.....	41
Figura 7 – Rede de ideias do cluster 0.....	43
Figura 8 – Rede de ideias do cluster 1.....	44
Figura 9 – Rede de ideias do cluster 2.....	45
Figura 10 – Rede de ideias do cluster 3.....	46
Figura 11 – Rede de ideias do cluster 4.....	47
Figura 12 – Gráfico de barra de Ideias de negócios por Ano.....	48
Figura 13 – Nuvem de palavras referente ao cluster 0.....	50
Figura 14 – Nuvem de palavras referente ao cluster 1.....	50
Figura 15 – Nuvem de palavras referente ao cluster 2.....	51
Figura 16 – Nuvem de palavras referente ao cluster 3.....	51
Figura 17 – Nuvem de palavras referente ao cluster 4.....	52
Figura 18 – Nuvem de palavras unindo os cinco clusters.....	53
Figura 19 – Mapa de calor da correlação de termos normalizados por TF-IDF - 10 termos - todos os dados.....	54
Figura 20 – Mapa de calor da correlação de termos normalizados por TF-IDF - 10 termos - pandemia (2019-2020).....	55
Figura 21 – Mapa de calor da correlação de termos normalizados por TF-IDF - 10 termos - últimos 3 anos (2021-2023).....	55

LISTA DE QUADROS

Quadro 1 – Metodologia DSRM.....	31
----------------------------------	----

LISTA DE ABREVIATURAS E SIGLAS

AI	<i>Artificial Intelligence</i> (Inteligência Artificial)
API	<i>Application Programming Interface</i>
CSV	<i>Comma-Separated Values</i> (Valores Separados por Vírgula)
DSRM	Metodologia de Pesquisa em Ciência de Design
IT	<i>Information Technology</i> (Tecnologia da Informação)
IVF	<i>In Vitro Fertilization</i> (Fertilização in Vitro)
IOT	<i>Internet of Things</i> (Internet das Coisas)
K-MEANS	<i>K-Means Clustering</i> (Agrupamento K-Means)
PCA	<i>Principal Component Analysis</i> (Análise de Componentes Principais)
TF-IDF	<i>Term Frequency-Inverse Document Frequency</i>
TIC	Tecnologias da Informação e Comunicação
LLMs	<i>Large Language Models</i> (Grande Modelo de Linguagem)

SUMÁRIO

1 INTRODUÇÃO	13
1.1 PROBLEMÁTICA	14
1.2 OBJETIVOS	14
1.2.1 Objetivo Geral	14
1.2.2 Objetivos Específicos	15
1.4 ESTRUTURA DO TEXTO	16
2 FUNDAMENTAÇÃO TEÓRICA	17
2.1 CONCEITOS DE INOVAÇÃO	17
2.1.1 História e evolução do conceito de inovação	17
2.1.2 Definição de inovação e sua importância no contexto empresarial	18
2.1.3 Dimensões da inovação	18
2.1.4 Tipos de inovação	18
2.1.5 Principais desafios na implementação da inovação	19
2.2 GESTÃO DE IDEIAS E INOVAÇÃO	19
2.2.1 Coleta, Avaliação, Seleção e Implementação de Ideias	20
2.2.2 Sistemas e Modelos de Gestão de Ideias	21
2.3 A ERA DA INFORMAÇÃO E A NECESSIDADE DE VISUALIZAÇÃO	22
2.3.1 Necessidade de Visualização	22
2.3.2 Tendências Futuras em Visualização de Dados	23
2.4 VISUALIZAÇÃO DE DADOS COM PYTHON	23
2.4.1 Bibliotecas e Benefícios da Visualização de Dados em Python	23
2.5 O PAPEL DA VISUALIZAÇÃO NA GERAÇÃO DE INSIGHTS	24
2.5.1 Transformação de Dados em Informação	24
2.5.2 Facilitando a Tomada de Decisão	24
2.5.3 Impulsionando a Inovação através da Visualização	25
2.6 ANÁLISE DE AGRUPAMENTO	26
2.6.1 Métodos e Técnicas	26
2.7 SITE "COOL BUSINESS IDEAS"	27
3 METODOLOGIA	29
3.1 DEFINIÇÃO DE PESQUISA	29
3.2 METODOLOGIA CIENTÍFICA E MÉTODO	29
3.3 TIPO DE PESQUISA	30
3.4 METODOLOGIA DSRM	30
3.4.1 População e Amostra	32
3.4.2 Coleta de Dados e Procedimentos	33
3.4.3 Análise de dados	34
3.4.4 Limitações da pesquisa	35
3.4.5 Resultados esperados	35
4 RESULTADOS E DISCUSSÕES	37
4.1 LIMPEZA E PREPARAÇÃO DOS DADOS	37
4.2 VETORIZAÇÃO E PRÉ-PROCESSAMENTO	38

4.3 REDUÇÃO DE DIMENSIONALIDADE E CLUSTERING	39
4.4 ANÁLISE EXPLORATÓRIA	42
4.4.1 Análise Detalhada dos clusters de Ideias de Negócios	42
4.4.2 Análise Temporal das Ideias de Negócios e Confirmação de Ideias Antigas	47
4.4.3 Análise de Palavras-chave e Tendências Temáticas	49
4.4.4 Análise de Palavras-chave e Mapas de Calor	53
4.4.5 Insights Extraídos das Visualizações	56
4.4.6 Implicações para a Inovação e Estratégia de Negócios	57
4.4.7 Limitações do Estudo e Considerações Futuras	58
5 CONSIDERAÇÕES FINAIS	60

1 INTRODUÇÃO

Em um cenário contemporâneo caracterizado pela geração contínua e maciça de dados, a capacidade de interpretar e compreender essas informações tornou-se essencial. Conforme argumentado por Padilla *et al.* (2018), a visualização de dados desempenha um papel crucial nesse contexto, sendo uma representação gráfica projetada para facilitar a compreensão intuitiva e acessível desses dados. Alhadad (2018) amplia ainda mais essa visão, enfatizando que a visualização não é apenas um meio de comunicação, mas também uma ferramenta poderosa para melhorar a percepção do público em relação a informações complexas.

O domínio da visualização de informações, como observado por Eberhard (2023), transcende as fronteiras de várias disciplinas e setores. Essa prática multifacetada desempenha um papel fundamental, desde concentrar a atenção até facilitar o compartilhamento do conhecimento e identificar tendências e padrões subjacentes nos dados. A amplitude de aplicação da visualização amplifica sua relevância, especialmente no contexto da inovação (WANG *et al.*, 2017).

No ambiente corporativo, a inovação é mais do que uma mera ferramenta competitiva; é o motor do progresso econômico, como afirmado por Hansen (2022). Baregheh, Rowley e Sambrook (2009) ecoam essa ideia, definindo inovação como a alquimia de transformar ideias em soluções tangíveis que agregam valor discernível ao mercado.

Entretanto, alimentar essa engrenagem de inovação requer uma gestão de ideias sólida e eficaz. Plataformas que agregam e apresentam ideias, como o "coolbusinessideas.com," desempenham um papel importante nesse contexto. No entanto, como apontam pesquisadores como Martinez-Torres e Olmedilla (2016), Gabriel *et al.* (2016) e Dziallas e Blind (2019), um desafio formidável para essas plataformas é o vasto oceano de informações não estruturadas que geram e que exigem análise metódica.

Este trabalho se propõe a explorar a interseção entre visualização de dados, inovação e gestão de ideias, destacando como a visualização pode ser uma ferramenta crítica para o processo de inovação e como abordagens eficazes de gestão de ideias podem ser aprimoradas por meio da análise de dados visualizados. A compreensão dessa relação é fundamental em um contexto empresarial cada vez mais orientado por dados e inovação.

1.1 PROBLEMÁTICA

A inovação desempenha um papel crucial no progresso corporativo e tecnológico, incentivando organizações a se adaptarem e permanecerem relevantes em um cenário que evolui continuamente. Em sua essência, a gestão eficaz de ideias é o cerne desse processo. Estudos como os de Björk, Boccardelli e Magnusson (2010) e Vandenbosch, Saatcioglu e Fay (2006) destacam que uma gestão eficiente de ideias serve como o catalisador fundamental para a identificação e materialização de oportunidades viáveis de inovação.

Contudo, no atual cenário digital, embora as ferramentas contemporâneas tenham oferecido inúmeras vantagens, também introduziram desafios sem precedentes. Plataformas como o "coolbusinessIdeas.com," enquanto importantes na orientação de ideias inovadoras, também contribuem para uma inundação de dados. A enxurrada de informações, frequentemente no formato textual, transformou a tarefa de selecionar e avaliar ideias em um verdadeiro desafio hercúleo (HOORNAERT *et al.*, 2017).

Dentro desse contexto, surge uma questão fundamental: Como as técnicas de visualização de dados podem ser efetivamente empregadas para discernir tendências, identificar padrões e extrair *insights* cruciais das ideias apresentadas em plataformas como o "coolbusinessIdeas.com"?

1.2 OBJETIVOS

Para melhor entendimento do trabalho apresentado, seus objetivos foram divididos entre objetivos gerais e específicos.

1.2.1 Objetivo Geral

Explorar as técnicas de visualização de dados para identificar tendências, padrões e *insights* nas ideias apresentadas na plataforma "coolbusinessIdeas.com", facilitando a visualização de ideias inovadoras.

1.2.2 Objetivos Específicos

Para alcançar o objetivo geral, os seguintes objetivos específicos são propostos:

- Coletar ideias apresentadas na plataforma "coolbusinessIdeas.com";
- Elaborar uma lista das ferramentas e técnicas de visualização de dados disponíveis, avaliando-as em termos de sua adequação para análise de ideias de negócios tecnológicos, com o objetivo de selecionar as mais eficazes para o estudo.
- Sintetizar os *insights* obtidos das visualizações de dados em um conjunto de recomendações estratégicas para gestão de ideias e inovação em tecnologia, com base nas tendências e padrões identificados nas ideias coletadas.

1.3 JUSTIFICATIVA

No dinâmico cenário de negócios de hoje, a inovação é a chave para não apenas sobreviver, mas prosperar. Como enfatizado por Baregheh, Rowley e Sambrook (2009), a inovação representa a habilidade de transformar ideias em soluções que conferem um valor excepcional ao mercado. A gestão eficaz de ideias emerge como o alicerce fundamental sobre o qual essa inovação é construída, com o potencial não apenas de impulsionar o crescimento econômico, como apontam Mahroum e Al-Saleh (2013), mas também de manter as organizações relevantes e ágeis em um ambiente empresarial em constante evolução.

No entanto, a gestão de ideias não está isenta de desafios complexos, como ressaltado por Martinez-Torres e Olmedilla (2016). Esses desafios incluem o crescente volume de dados, a recorrência de ideias redundantes e a pressão por avaliação e implementação ágeis. A complexidade intrínseca desses obstáculos amplifica a necessidade de investigar e aplicar técnicas eficazes, como a visualização de dados, a fim de extrair *insights* valiosos de plataformas de ideias.

A visualização de dados, como sugerido por Norris (2022), emerge como uma ferramenta essencial que pode iluminar o caminho através deste intrincado labirinto de informações. Ao otimizar o processo de inovação, a visualização de dados não apenas facilita a compreensão rápida e a identificação de padrões nas ideias, mas também transforma informações complexas em *insights* acionáveis. Isso, por sua

vez, acelera o ciclo de inovação e concede às organizações uma vantagem competitiva vital em um mercado caracterizado pela concorrência incessante e pelas mudanças constantes.

1.4 ESTRUTURA DO TEXTO

O conteúdo desta monografia foi estruturado em cinco capítulos a fim de facilitar o entendimento acerca do que foi desenvolvido, sendo eles:

A introdução, que apresenta a contextualização do trabalho, a motivação, a problemática, os objetivos e a justificativa do estudo.

No segundo capítulo, aborda-se o referencial teórico, discutindo conceitos fundamentais sobre visualização de dados, gestão de ideias e inovação.

A metodologia adotada é descrita no terceiro capítulo, detalhando o método de pesquisa escolhido, as ferramentas de visualização utilizadas e o processo seguido para a coleta e análise de dados.

No quarto capítulo, apresentam-se os resultados obtidos, incluindo visualizações e análises, assim como uma discussão sobre os insights e tendências identificadas.

Por fim, o quinto capítulo traz as considerações finais sobre o trabalho, refletindo sobre os resultados alcançados, as contribuições do estudo e possíveis caminhos para trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Nesta seção, são apresentados alguns conceitos e terminologias que serão utilizados ao longo deste trabalho.

2.1 CONCEITOS DE INOVAÇÃO

Inovação é um termo amplamente utilizado no mundo dos negócios e da academia, servindo como um pilar fundamental para o progresso e desenvolvimento contínuo. Ao longo das décadas, sua definição e aplicação têm evoluído, refletindo as mudanças nas necessidades da sociedade e nas dinâmicas do mercado.

No contexto empresarial, a inovação é frequentemente vista como a chave para ganhar vantagem competitiva, impulsionar o crescimento e responder eficazmente às demandas em constante mudança dos consumidores (BAREGHEH *et al.*, 2009).

No entanto, compreender verdadeiramente o que é inovação e como ela se manifesta requer uma exploração mais profunda de seus diversos aspectos e dimensões. Esta seção busca esclarecer e expandir o conceito de inovação, explorando sua história, dimensões, tipos e os desafios associados à sua implementação (CARBONE *et al.*, 2012).

2.1.1 História e evolução do conceito de inovação

A inovação, ao longo do tempo, tem sido a força motriz por trás do avanço da sociedade e da economia. A compreensão do conceito evoluiu de simplesmente introduzir novidades para um entendimento mais abrangente que engloba a implementação bem-sucedida dessas novidades no mercado (BAREGHEH *et al.*, 2009).

A história da inovação é repleta de exemplos de ideias que transformaram setores, da revolução industrial à era digital. Pense, por exemplo, nas mudanças trazidas pela invenção da eletricidade, pela introdução dos computadores pessoais, ou pela ascensão da *internet*. Estas inovações não só transformaram setores inteiros, mas também moldaram comportamentos e formas de vida (GABRIEL *et al.*, 2016; HANSEN, 2022). Esta evolução contínua reforça a ideia de que a inovação

não é apenas sobre invenção, mas sobre a capacidade de adaptar e implementar novas soluções em um contexto real, gerando valor e impacto.

2.1.2 Definição de inovação e sua importância no contexto empresarial

No contexto empresarial, a inovação refere-se à capacidade de uma organização introduzir algo novo ou diferente que gera valor, seja para seus *stakeholders*, clientes ou para a sociedade em geral (BAREGHEH *et al.*, 2009). Esta capacidade não se limita apenas à introdução de novos produtos ou serviços, mas também abrange a implementação de novos processos, estratégias e formas de pensar (BREM; VOIGT, 2007).

A importância da inovação no contexto empresarial é inegável, pois é através dela que as empresas conseguem criar vantagens competitivas, diferenciar-se em mercados saturados e responder de forma ágil e eficaz às constantes mudanças do ambiente de negócios (MAHROUM; AL-SALEH, 2013).

2.1.3 Dimensões da inovação

A inovação não se restringe a uma única dimensão. Ela se manifesta de várias formas, desde inovações tecnológicas até mudanças organizacionais e processuais. A inovação tecnológica, por exemplo, refere-se à introdução de novas tecnologias ou aprimoramentos significativos nas existentes.

Por outro lado, inovações organizacionais e processuais focam na forma como as empresas operam e entregam valor aos seus clientes (BJÖRK; BOCCARDELLI; MAGNUSSON, 2010). É essencial reconhecer e valorizar todas essas dimensões, pois cada uma delas desempenha um papel crucial na construção de empresas resilientes e competitivas em um mercado globalizado (DZIALLAS; BLIND, 2019).

2.1.4 Tipos de inovação

A inovação pode ser categorizada com base em sua natureza e impacto. A inovação incremental refere-se a melhorias contínuas em produtos ou processos já

existentes, enquanto a inovação radical envolve a criação de soluções completamente novas, com o potencial de transformar mercados (DZIALLAS; BLIND, 2019). Além dessas, tem-se a inovação de produto, focada na introdução ou aprimoramento de produtos, e a inovação de processo, que se concentra em novos ou melhorados métodos de produção ou entrega (BJÖRK; BOCCARDELLI; MAGNUSSON, 2010).

2.1.5 Principais desafios na implementação da inovação

Apesar de seu potencial transformador, a inovação não está isenta de desafios. Implementar novas ideias em um contexto empresarial pode enfrentar resistências internas, como a aversão ao risco e a resistência a mudanças por parte dos colaboradores (BAKKER, 2010).

Além disso, há barreiras regulatórias que podem dificultar a entrada de novos produtos ou serviços no mercado, especialmente em setores altamente regulamentados como saúde ou finanças (CARBONE *et al.*, 2012). A velocidade da mudança tecnológica, juntamente com as crescentes expectativas dos clientes, exige das empresas uma capacidade de adaptação e aprendizado contínuos.

Portanto, para que a inovação seja bem-sucedida, é essencial uma abordagem estratégica, focada não apenas na geração de ideias, mas também na sua implementação eficaz, na formação contínua dos colaboradores e na superação dos desafios inerentes ao processo.

2.2 GESTÃO DE IDEIAS E INOVAÇÃO

No cenário empresarial contemporâneo, marcado por mudanças rápidas e evoluções tecnológicas constantes, a habilidade de gerar, avaliar e concretizar ideias inovadoras é essencial para manter-se relevante e competitivo. Sites como CoolBusinessIdeas.com exemplificam essa necessidade, atuando como hubs centrais para disseminação, avaliação e refinamento de ideias de negócios.

A gestão de ideias surge como uma abordagem estratégica que busca transformar *insights* iniciais em inovações tangíveis. Este processo, contextualizado no mundo dos negócios, envolve etapas de coleta, avaliação, seleção e implementação de ideias, todas com o propósito primordial de criar valor e

estabelecer uma vantagem competitiva (VANDENBOSCH; SAATCIOGLU, 2006). Mais do que apenas um método, é o alicerce para a introdução de novos produtos, serviços ou processos que atendem demandas de mercado ou solucionam problemas específicos.

Bakker (2010) e Gibson e Skarzynski (2008) enfatizam a importância de compreender os processos criativos e a dinâmica de geração de ideias para alcançar a inovação sustentada. Ao incorporar efetivamente a gestão de ideias, as organizações não apenas potencializam sua capacidade inovadora, mas também cultivam a criatividade e o engajamento de seus *stakeholders*, estabelecendo um ambiente fértil para a contínua criação de valor.

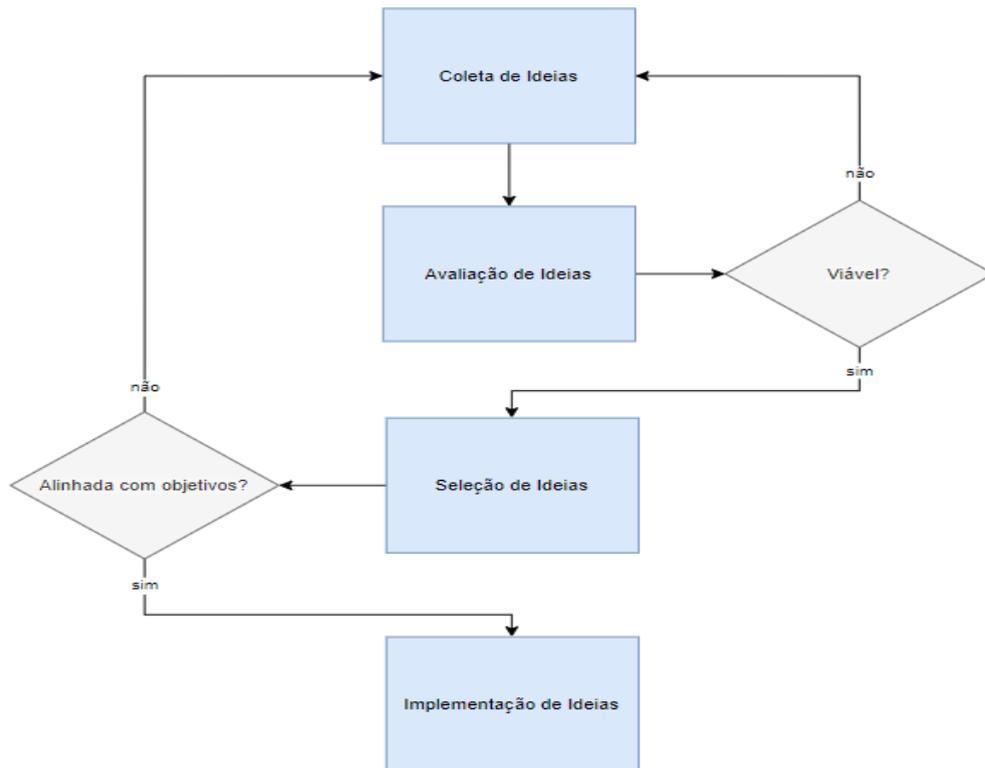
2.2.1 Coleta, Avaliação, Seleção e Implementação de Ideias

A coleta de ideias é o ponto de partida na gestão eficaz de inovações. As organizações modernas reconhecem o potencial que reside em diferentes *stakeholders* e, por isso, incentivam ativamente seus funcionários, clientes e até o público em geral a compartilhar suas ideias e *insights*. Métodos variados, como *brainstorming*, *workshops* e plataformas digitais, são empregados nesse processo de coleta (LEE *et al.*, 2018).

Após a coleta, é imperativo avaliar essas ideias. Schumpeter (1949) e Gaynor (2001) salientam que ideias bem gerenciadas e executadas têm o potencial de se transformar em inovações valiosas. Essa avaliação é feita com base em critérios como viabilidade, alinhamento estratégico e potencial de mercado. Adams, Bessant e Phelps (2006) destacam a importância da filtragem e seleção adequadas, considerando-as fases críticas. As ideias que se alinham com a visão e objetivos da organização são, então, selecionadas para futura implementação (GAMA; FRISHAMMAR; PARIDA, 2019).

Uma vez selecionadas, as ideias entram em uma fase intensiva de transformação. Aqui, elas evoluem de meros conceitos para soluções tangíveis que podem ser produtos, serviços ou processos. Koen *et al.* (2002) discutem a intrincada natureza do processo de inovação, sublinhando seu caráter iterativo. Esta fase é repleta de desafios, como a resistência à mudança, alocação de recursos e a necessidade de manter o alinhamento estratégico. O fluxograma na Figura 1 apresenta visualmente as etapas citadas.

Figura 1 – Fluxograma de Processo de Gestão de Ideias: Coleta, Avaliação, Seleção e Implementação.



Fonte: elaborado pelo autor

2.2.2 Sistemas e Modelos de Gestão de Ideias

Para otimizar a gestão de ideias, surgiram Sistemas de Gestão de Ideias, projetados especificamente para gerenciar ideias coletadas dentro das organizações (POVEDA; WESTERSKI; IGLESIAS, 2012). Tais sistemas, além de coletar, oferecem ferramentas que auxiliam na avaliação, seleção e implementação das ideias.

Empresas reconhecidas mundialmente, como *Dell*®, *Starbucks*® e *Cisco*®, têm investido em aplicativos de Gestão de Ideias para impulsionar a inovação e melhorar seus produtos e serviços. Diversos modelos foram propostos ao longo dos anos para guiar a Gestão de Ideias, desde abordagens mais tradicionais, como a de Cooper (1990), até modelos mais contemporâneos, como o de Flynn *et al.* (2003), refletindo a evolução e a diversidade no campo da gestão de ideias.

2.3 A ERA DA INFORMAÇÃO E A NECESSIDADE DE VISUALIZAÇÃO

A transição para a era digital trouxe consigo uma revolução sem precedentes em termos de produção e acesso à informação. A conectividade global e a ascensão da Internet transformaram a maneira como comunica-se, trabalha-se e relaciona-se, gerando uma quantidade massiva de dados a cada instante (EBERHARD, 2023). Esse crescimento exponencial de informações é impulsionado não apenas por interações em mídias sociais e transações online, mas também por avanços como a Internet das Coisas (IoT) e a análise de *big data* (POLHEMUS *et al.*, 2022; HILBERT, 2016).

Com o aumento constante na produção de dados, surgem desafios significativos. A sobrecarga de informações, ou "*infoxication*", pode dificultar a identificação de informações relevantes e a tomada de decisões informadas (CHEN *et al.*, 2022). Além disso, a variedade e velocidade com que os dados são gerados exigem ferramentas e técnicas mais avançadas para seu processamento e análise (BORGES, 2015).

2.3.1 Necessidade de Visualização

Frente a esse cenário, a visualização de dados emerge como uma solução fundamental. Transformando conjuntos complexos de dados em representações gráficas intuitivas, a visualização facilita a compreensão e interpretação das informações (PADILLA *et al.*, 2018). Esta técnica torna-se essencial não apenas para identificar tendências e padrões, mas também para guiar políticas públicas, estratégias empresariais e inovações (MUNZNER, 2014).

As tecnologias emergentes, como realidade virtual e aumentada, estão elevando a visualização a novos patamares, permitindo experiências mais imersivas e *insights* mais detalhados (KOVACOVA *et al.*, 2022; WESTERSKI & IGLESIAS, 2011).

2.3.2 Tendências Futuras em Visualização de Dados

O campo da visualização de dados está em constante evolução. Com a integração da Inteligência Artificial e *Machine Learning*, novas ferramentas estão surgindo, capazes de oferecer visualizações mais dinâmicas e personalizadas (WONG *et al.*, 2010). Além disso, a visualização imersiva, através de realidade virtual e aumentada, promete revolucionar áreas como medicina, educação e *design* urbano, proporcionando experiências mais ricas e contextualizadas (KOVACOVA *et al.*, 2022).

2.4 VISUALIZAÇÃO DE DADOS COM PYTHON

Python, emergindo como uma das linguagens de programação mais populares, é amplamente reconhecida por sua simplicidade e versatilidade. Particularmente no domínio da ciência de dados. A linguagem destaca-se pela sua rica coleção de bibliotecas dedicadas à visualização de dados, tornando a interpretação e apresentação de informações complexas mais acessível e intuitiva (Munzner, 2014).

2.4.1 Bibliotecas e Benefícios da Visualização de Dados em Python

Python possui várias bibliotecas que são essenciais para visualização de dados, com Matplotlib, Seaborn e Plotly sendo algumas das mais destacadas. A Matplotlib é amplamente considerada a fundação da visualização de dados em Python, fornecendo uma interface versátil para criar gráficos estáticos, animados e interativos. Seaborn, construído em cima da Matplotlib, simplifica a criação de visualizações mais atraentes e informativas. Plotly, por outro lado, é conhecida por sua capacidade interativa avançada, permitindo a criação de visualizações dinâmicas adequadas para aplicações *web* (MUNZNER, 2014; CHEN *et al.*, 2022).

A capacidade de visualizar dados usando essas bibliotecas não apenas facilita a compreensão de complexidades nos dados, mas também permite a comunicação eficaz de *insights*. A possibilidade de criar visualizações interativas, em particular, proporciona uma experiência mais envolvente, permitindo aos usuários

explorar e interagir com os dados de forma mais intuitiva. Além disso, a vasta comunidade Python garante uma constante atualização e melhoria dessas ferramentas, garantindo que permaneçam relevantes e eficientes no tratamento das crescentes demandas da análise de dados (POLHEMUS *et al.*, 2022).

2.5 O PAPEL DA VISUALIZAÇÃO NA GERAÇÃO DE *INSIGHTS*

A visualização de dados não é apenas uma representação estética, é uma ferramenta de comunicação poderosa que facilita a compreensão e a interpretação de conjuntos de dados complexos. Ao transformar dados brutos em representações visuais intuitivas, a visualização ajuda a descobrir *insights* ocultos e a informar a tomada de decisões (PADILLA *et al.*, 2018).

Para plataformas como CoolBusinessIdeas.com, que visam identificar e apresentar as ideias de negócios mais inovadoras do momento, a visualização é vital para destacar tendências emergentes e oferecer *insights* claros.

2.5.1 Transformação de Dados em Informação

Os dados, em sua forma bruta, podem ser opacos e difíceis de interpretar. No entanto, quando visualizados, eles ganham vida e contam uma história. Gráficos, diagramas, mapas de calor e outras representações visuais têm o poder de revelar padrões, mostrar tendências emergentes e destacar anomalias que podem não ser imediatamente aparentes em tabelas ou relatórios densos (ALHADAD, 2018). Por exemplo, um gráfico de linha pode mostrar a trajetória de crescimento de uma empresa ao longo do tempo, enquanto um gráfico de dispersão pode revelar correlações entre variáveis que, de outra forma, poderiam passar despercebidas (WESTERSKI; IGLESIAS, 2011).

2.5.2 Facilitando a Tomada de Decisão

Num mundo onde a tomada de decisões baseadas em dados se tornou a norma, a visualização serve como uma ponte entre dados brutos e *insights* acionáveis. Ao representar informações de forma clara e acessível, as visualizações

permitem que os tomadores de decisão assimilem rapidamente informações complexas, levando a decisões mais informadas e, em última instância, a melhores resultados (EBERHARD, 2023).

Além disso, as modernas ferramentas de visualização de dados são interativas, permitindo que os usuários manipulem os dados, realizem análises *ad-hoc* e explorem diferentes cenários, tudo isso em tempo real. Essa capacidade de interatividade amplifica a utilidade das visualizações, tornando-as não apenas instrumentos de observação, mas também ferramentas de exploração e descoberta (HILBERT, 2016).

2.5.3 Impulsionando a Inovação através da Visualização

Em diversos setores, a inovação é frequentemente alimentada por *insights* extraídos de dados. Nesse contexto, a visualização de dados atua como uma alavanca, amplificando a capacidade das organizações de identificar oportunidades e desafios emergentes. No campo da saúde, visualizações detalhadas podem ajudar os pesquisadores a identificar padrões em dados de pacientes, levando potencialmente a novos tratamentos ou métodos de prevenção (HILBERT, 2016). No mundo dos negócios, a descoberta de tendências de consumo ou de eficiência operacional através da visualização pode revelar oportunidades de mercado previamente negligenciadas (WESTERSKI; IGLESIAS, 2011).

Conforme destacado por CHEN *et al.* (2022), a visualização desempenha um papel crítico na descoberta de conhecimento em grandes conjuntos de dados. Ao revelar relações, correlações e anomalias, a visualização capacita as organizações a agir com base em *insights* concretos.

Em um ambiente onde a capacidade de inovar rapidamente define líderes de mercado, a visualização de dados não é apenas uma ferramenta técnica, mas uma competência estratégica. Ela permite que as organizações capitalizem os *insights* derivados de seus dados, conduzindo a decisões mais informadas e inovações transformadoras (EBERHARD, 2023).

2.6 ANÁLISE DE AGRUPAMENTO

A Análise de Agrupamento, também denominada por diferentes nomenclaturas, como Taxonomia Numérica, *Data clustering*, e Análise de *clusters* (ANDRIENKO; ANDRIENKO, 2009), é uma técnica empregada na exploração e análise de dados. O principal propósito desta técnica é descobrir e interpretar grupos de objetos que possuem propriedades e/ou comportamentos semelhantes (ANDRIENKO; ANDRIENKO, 2009; KONCHADY, 2006).

A técnica de análise de agrupamento é classificada como uma forma de aprendizado não supervisionado, uma vez que não existe uma classificação prévia dos dados para o subsequente agrupamento (KONCHADY, 2006). Esta abordagem é particularmente útil na análise de vastos volumes de dados, pois permite ao analista focar em grupos de objetos ao invés de objetos individuais.

A análise de agrupamento no contexto do aprendizado de máquina envolve o uso de algoritmos para processar e analisar grandes volumes de dados, agrupando-os com base em similaridades identificadas automaticamente. Esses métodos são essenciais para identificar padrões e *insights* em grandes conjuntos de dados, uma capacidade cada vez mais vital em diversas aplicações práticas e científicas. A evolução contínua desses métodos e técnicas evidencia o papel crescente do aprendizado de máquina no manejo eficiente e eficaz de grandes volumes de dados (Zhang *et al.*, 2023; Khan & Siddiqui, 2023).

2.6.1 Métodos e Técnicas

A análise de agrupamento pode ser realizada usando diversos métodos, cada um com suas características e nuances. O método hierárquico é uma dessas abordagens que constrói uma hierarquia de dados, geralmente visualizada como um dendrograma. Existem duas abordagens principais dentro deste método: aglomerativa e divisiva. A primeira começa com cada objeto como um *cluster* individual e combina grupos até todos os objetos estarem em um único *cluster*. Em contraste, a abordagem divisiva começa com todos os objetos em um único *cluster* e os divide até que cada objeto esteja em seu próprio *cluster* (ANDRIENKO; ANDRIENKO, 2009; JAIN; DUBES, 1988).

Por outro lado, os métodos baseados em particionamento, como o *k-means*, focam em determinar o número de *clusters* a serem formados de antemão. O *k-means* é um dos algoritmos mais populares nesta categoria e tem sido amplamente utilizado devido à sua eficiência em minimizar as distâncias entre os objetos e o centróide de seu *cluster* (JAIN; MURTY; FLYNN, 1999).

Antes da aplicação desses métodos, técnicas como TF-IDF e PCA podem ser implementadas para preparar e otimizar os dados. O TF-IDF é uma medida estatística usada para avaliar a importância de uma palavra em um documento em relação a uma coleção ou *corpus*, destacando palavras-chave significativas (RAMOS, 2003). Isso é especialmente relevante em conjuntos de dados textuais onde a identificação de termos únicos é necessária para discernir o conteúdo de cada documento. Uma vez processados pelo TF-IDF, os dados muitas vezes requerem uma redução de dimensionalidade para facilitar a análise de agrupamento. A PCA é uma técnica de redução de dimensionalidade que transforma os dados originais em um número menor de variáveis independentes chamadas componentes principais, preservando a maior parte da informação (JOLLIFFE, 2002). Esses componentes principais servem como a base para o agrupamento subsequente, seja hierárquico ou particional, e são cruciais para uma visualização eficaz e interpretação dos dados.

2.7 SITE "COOL BUSINESS IDEAS"

CoolBusinessIdeas.com é uma plataforma dedicada à coleta de novas e promissoras ideias e oportunidades de negócios de todo o mundo. No ambiente digital atual, plataformas online como esta, desempenham um papel vital, agindo como hubs centrais para a disseminação, avaliação e aprimoramento de ideias de negócios (SALDIVAR *et al.*, 2016; WESTERSKI; IGLESIAS, 2011).

Seu objetivo é informar seus leitores sobre tendências emergentes no mundo dos negócios, ajudando-os a ganhar uma vantagem sobre a concorrência. Assim, empresas e empreendedores têm a possibilidade de recorrer à plataforma para identificar oportunidades de inovação, bem como para confirmar e refinar suas próprias ideias de negócios (MARTINEZ-TORRES; OLMEDILLA, 2016; LEE *et al.*, 2018).

A plataforma destaca-se por sua abordagem proativa de rastrear novas ideias de negócios, aproveitando uma ampla rede de informantes e pesquisa própria. Em essência, CoolBusinessIdeas.com serve como um recurso valioso para empreendedores e profissionais de negócios, fornecendo *insights* sobre as ideias de negócios mais inovadoras do momento. O site exemplifica a evolução das estratégias de inovação aberta, onde a colaboração e a coleta de *insights* de uma ampla comunidade tornam-se fundamentais para o processo de inovação (CARBONE *et al.*, 2012; GAMA, FRISHAMMAR; PARIDA, 2019).

3 METODOLOGIA

Este capítulo tem como objetivo fornecer uma descrição abrangente dos métodos e técnicas empregados na condução da pesquisa.

3.1 DEFINIÇÃO DE PESQUISA

O trabalho apresentado, sob o ponto de vista de sua natureza é caracterizado por uma pesquisa aplicada uma vez que objetiva gerar conhecimentos para aplicação prática, direcionados à solução de problemas específicos.

Segundo Gil (2008, p. 17), o termo pesquisa é definido como o:

[...] procedimento racional e sistemático que tem como objetivo proporcionar respostas aos problemas que são propostos. A pesquisa desenvolve-se por um processo constituído de várias fases, desde a formulação do problema até a apresentação e discussão dos resultados.

Planejar passo a passo os procedimentos utilizados para o desenvolvimento da pesquisa abrange a primeira fase da pesquisa científica, envolvendo desde a escolha do tema até a operacionalização dos métodos.

3.2 METODOLOGIA CIENTÍFICA E MÉTODO

Segundo Tartuce (2006) a metodologia científica versa sobre método e ciência. O termo método está relacionado ao caminho que será traçado para alcançar um fim e o termo ciência está relacionado ao conjunto de conhecimentos precisos e ordenados em relação a determinado domínio do saber. Segundo o autor, o termo metodologia se refere ao estudo do método, objetivando estabelecer regras e procedimentos para realizar uma pesquisa. Assim, a metodologia científica é o estudo sistemático e lógico dos métodos utilizados nas ciências, envolvendo seus fundamentos, confiabilidade e relação com as teorias científicas (TARTUCE, 2006).

3.3 TIPO DE PESQUISA

O estudo realizado é de natureza exploratória, cujo principal objetivo é identificar e compreender as tendências emergentes no domínio da tecnologia, particularmente em relação às ideias de negócios propostas no site CoolBusinessIdeas.com.

A pesquisa exploratória é frequentemente usada quando se busca obter uma compreensão inicial sobre um tópico ou quando o campo de estudo é relativamente desconhecido. Ela permite que os pesquisadores se familiarizem com detalhes específicos, conceitos e variáveis intrínsecas ao tema (GIL, 2008).

Neste contexto, a pesquisa não se limita apenas a identificar as ideias de negócios, mas também a agrupá-las com base em sua similaridade de conteúdo. Esse agrupamento facilita a identificação de temas ou tópicos predominantes, fornecendo *insights* valiosos sobre áreas potenciais de inovação e oportunidades de mercado no setor tecnológico.

3.4 METODOLOGIA DSRM

Para orientar o desenvolvimento da pesquisa, foi adotada a Metodologia de Pesquisa em Ciência de Design (*Design Science Research Methodology*). A DSRM é uma abordagem amplamente utilizada em disciplinas como Ciência da Computação, Sistemas de Informação e Engenharia de *Software*.

A metodologia DSRM tem como objetivo principal projetar e desenvolver soluções para problemas práticos, frequentemente relacionados a sistemas de informação e inovação. Ela enfatiza que os pesquisadores não apenas estudam problemas, mas também contribuem para sua resolução, projetando soluções eficazes. Esta metodologia é particularmente reconhecida pela sua aplicabilidade na criação e avaliação de artefatos tecnológicos com o propósito de resolver problemas identificados e contribuir para o conhecimento científico (PEFFERS, K. *et al.*, 2007).

A abordagem DSRM segue um ciclo iterativo, que compreende as seguintes etapas:

- Identificação do Problema
- Definição dos Requisitos

- Design da Solução
- Desenvolvimento e Implementação
- Avaliação e Confirmação
- Comunicação dos Resultados
- Refinamento Iterativo (se necessário)

O Quadro 1 apresenta as etapas da metodologia preenchidas:

Quadro 1 – Metodologia DSRM

Identificação do problema	Exploração das tendências tecnológicas emergentes por meio da análise de ideias de negócios no site CoolBusinessIdeas.com, visando compreender o panorama atual e as oportunidades de inovação.
Definição dos requisitos	Definição de parâmetros para a análise exploratória, com o objetivo de obter <i>insights</i> iniciais sobre tendências em uma área de estudo em evolução.
Design da Solução	Agrupamento das ideias de negócios com base na similaridade do conteúdo através do emprego de técnicas de processamento de linguagem natural e aprendizado de máquina, facilitando a análise e interpretação dos dados.
Desenvolvimento e Implementação	Emprego de processamento de linguagem natural e aprendizado de máquina para processamento e análise detalhada dos dados, visando extrair padrões significativos e tendências das ideias de negócios coletadas.
Avaliação e Confirmação	Avaliação dos <i>clusters</i> formados para garantir coesão e relevância temática, complementada por visualizações para verificar qualitativamente as tendências observadas.
Comunicação dos Resultados	Apresentação dos <i>insights</i> obtidos e das interpretações derivadas, enfatizando as tendências predominantes e suas implicações estratégicas para a inovação e o desenvolvimento de negócios.
Refinamento Iterativo	Reciclagem do processo de pesquisa, se necessário, utilizando os resultados da análise de clusters e descobertas subsequentes para refinar e aprimorar o agrupamento das ideias de negócios.

Fonte: elaborado pelo autor

Essas etapas compõem uma estrutura sólida para a condução da pesquisa, desde a identificação do problema até a comunicação dos resultados. Para uma compreensão maior, detalha-se o passo a passo desta pesquisa nas próximas seções.

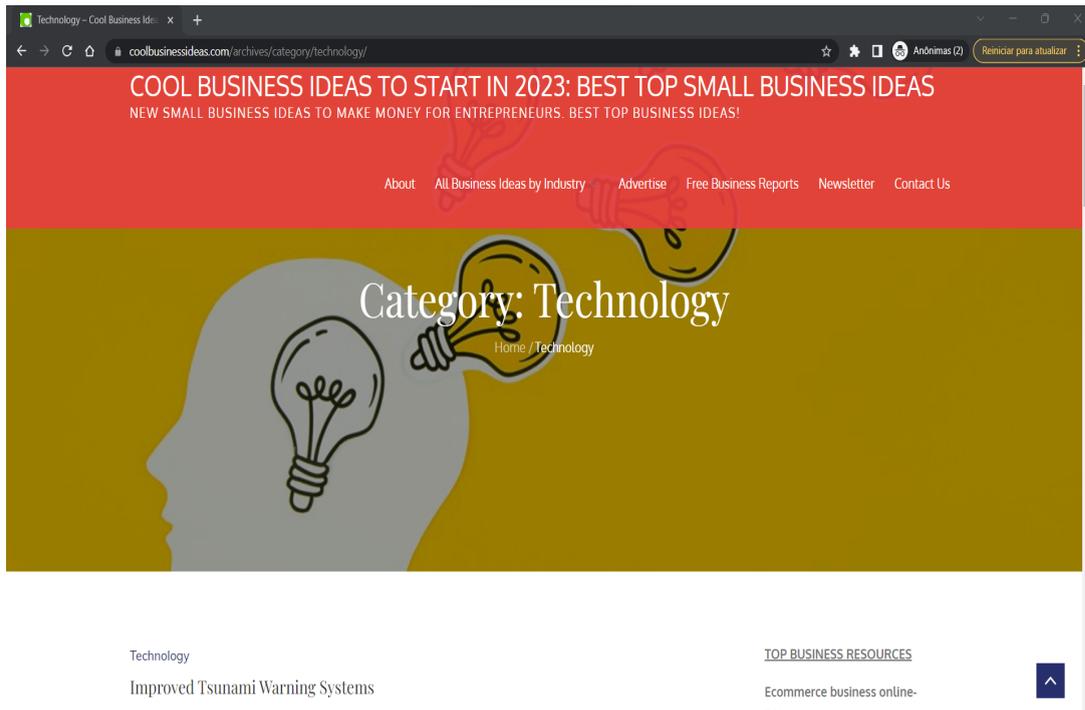
3.4.1 População e Amostra

O site CoolBusinessIdeas.com atua como um repositório global de ideias de negócios, abrangendo diversas categorias representando diferentes domínios e setores industriais. Cada ideia postada no site é considerada uma unidade individual dentro da população total de ideias de negócios disponíveis.

Para este estudo, concentrou-se na análise da categoria "Tecnologia", ilustrada na Figura 2, devido à sua relevância e potencial impacto na sociedade moderna e na economia digital. Essa categoria engloba inovações e tendências emergentes no domínio tecnológico.

A amostra selecionada para análise compreendeu 904 ideias de negócios da categoria "Tecnologia" postadas no período de 2004 a 2023. Essa seleção proporcionou uma visão abrangente das tendências e evoluções no setor tecnológico ao longo de duas décadas. Por meio desse recorte temporal, busca-se identificar padrões, temas recorrentes e possíveis previsões de direções futuras para inovações tecnológicas. Cada ideia de negócio na amostra foi representada por seu título e descrição, servindo como indicadores primários do conteúdo e contexto da inovação proposta.

Figura 2 – Página da categoria 'Tecnologia' no site CoolBusinessIdeas.com



Fonte: elaborado pelo autor

3.4.2 Coleta de Dados e Procedimentos

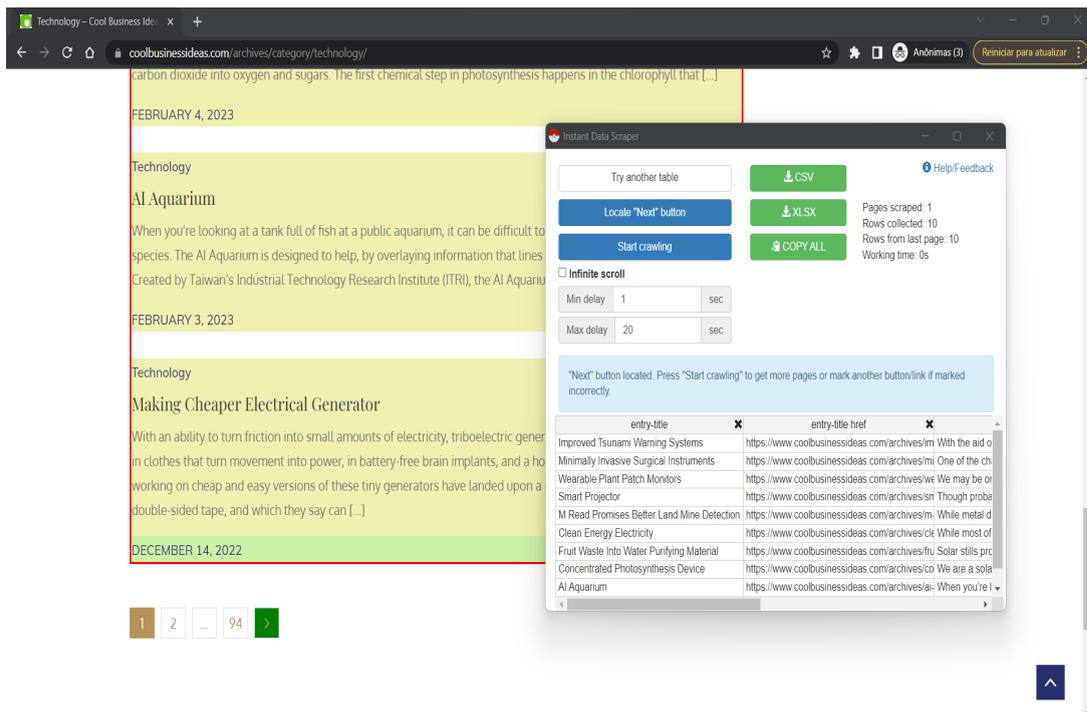
A metodologia adotada para a coleta de dados baseou-se na técnica de *web scraping*, que envolve a extração automatizada de informações de páginas da *web*. Dado que o site CoolBusinessIdeas.com não oferecia um conjunto de dados predefinido ou uma API para acessar suas informações, a técnica de *scraping* foi identificada como a abordagem mais viável para coletar as ideias de negócios necessárias para este estudo. É importante destacar que, para esta pesquisa, foram coletados apenas os resumos das ideias de negócios disponíveis no site, e não os textos completos. Essa escolha foi feita devido à natureza do site, que apresenta as ideias em forma de resumos concisos, oferecendo uma visão geral das inovações.

O processo de coleta de dados incluiu as seguintes etapas:

- Navegação e Seleção: Navegação manual pelo site CoolBusinessIdeas.com, com foco na categoria "Tecnologia";
- Uso do *Instant Data Scraper*: Utilização da extensão "*Instant Data Scraper*" do *Google Chrome* para automatizar a coleta de dados;

- Seleção de Colunas de Dados: Especificação das colunas de interesse para a coleta, incluindo os títulos e descrições das ideias de negócios;
- Configuração do Paginador: Configuração do paginador para automatizar a coleta de dados em múltiplas páginas;
- Início do Raspador de Dados (*Crawling*): Início do processo de coleta de dados usando a extensão "*Instant Data Scraper*";
- Download do CSV: Após a coleta de dados em todas as páginas necessárias, os dados foram baixados em um arquivo CSV;
- Exportação dos Dados: Exportação dos dados coletados em formato CSV para análise posterior, conforme ilustrado na Figura 3.

Figura 3 – Extração de dados com a extensão 'Instant Data Scraper'



Fonte: elaborado pelo autor

3.4.3 Análise de dados

A fase de análise de dados é crucial para extrair significados e *insights* do conjunto de informações coletadas. Este estudo adota uma abordagem

quantitativa, utilizando técnicas de aprendizado de máquina para agrupar e interpretar as ideias de negócios da categoria "Tecnologia".

Os procedimentos a seguir compõem a análise de dados:

- Vetorização do Conteúdo com TF-IDF: Transformação dos dados textuais em uma matriz de termos TF-IDF, permitindo a avaliação da importância relativa das palavras no corpus das ideias de negócios.
- Agrupamento com K-Means: Aplicação do algoritmo K-Means para agrupar as ideias de negócios com base em sua similaridade de conteúdo. O número de grupos (k) foi determinado durante a análise exploratória.
- Visualização de Dados e Redução de Dimensionalidade com PCA: Uso da técnica de Análise de Componentes Principais (PCA) para simplificar a complexidade dos dados e facilitar a visualização dos *clusters*.
- Interpretação Preliminar dos *clusters*: Análise inicial dos *clusters* para identificar tendências, similaridades e diferenças, fornecendo direcionamento para a discussão e apresentação dos resultados subsequentes.

3.4.4 Limitações da pesquisa

Este estudo enfrenta várias limitações, incluindo a concentração em uma única categoria, o uso de *web scraping* sujeito a inconsistências e a natureza dinâmica do site CoolBusinessIdeas.com, que pode afetar a qualidade dos dados coletados. Recomenda-se levar em conta essas limitações ao interpretar os resultados, considerando-os indicativos para futuras investigações que possam superar essas barreiras metodológicas.

3.4.5 Resultados esperados

Espera-se que esta pesquisa alcance os seguintes resultados:

- Identificação de *Insights* e Tendências nas Ideias de Negócios: Através da análise dos dados coletados, espera-se descobrir percepções e padrões emergentes que possam fornecer uma visão profunda sobre as tendências em ideias de negócios;

- Aprimoramento da Gestão de Ideias: Com os *insights* obtidos, busca-se proporcionar melhorias na forma como as ideias de negócios são gerenciadas e avaliadas;
- Recomendações de Ferramentas e Técnicas de Visualização: Identificar ferramentas e técnicas específicas que se mostrem mais eficazes para a visualização e análise das ideias coletadas, possibilitando uma análise mais intuitiva e profunda;

Além desses resultados diretos, a pesquisa tem o potencial de gerar contribuições acadêmicas, como a publicação de resultados, desenvolvimento de *software* ou protótipo e parcerias acadêmicas em áreas relacionadas à análise de dados e visualização.

4 RESULTADOS E DISCUSSÕES

Este capítulo apresenta os resultados da análise das ideias de negócios na categoria de tecnologia coletadas do site CoolBusinessIdeas.com. A análise busca discernir padrões e tendências.

Foram examinadas 904 ideias de negócios reunidas no período de 2004 a 2023. A abordagem adotada compreendeu a aplicação da técnica de vetorização TF-IDF e a utilização do método de agrupamento K-Means. Posteriormente, foi realizada a redução de dimensionalidade por meio de PCA para facilitar a visualização dos dados.

Os detalhes de cada etapa analítica serão apresentados, acompanhados de visualizações gráficas apropriadas, e será discutida a relevância dos achados em relação aos objetivos de pesquisa propostos.

4.1 LIMPEZA E PREPARAÇÃO DOS DADOS

Nesta fase do projeto, o foco foi aprimorar a qualidade dos dados para análise. Inicialmente, os dados foram carregados a partir do arquivo “coolbusinessideas.csv” utilizando a biblioteca *Pandas*. A limpeza e o pré-processamento dos dados envolveram várias etapas essenciais para garantir sua consistência e relevância.

Primeiramente, as colunas “*entry-title*” e “*entry-content*”, que contêm o título e o conteúdo textual das ideias de negócios, foram submetidas a um processo de normalização. Este processo incluiu a conversão do texto para letras minúsculas e a remoção de caracteres especiais, utilizando expressões regulares para garantir a uniformidade do texto. Essa normalização é crucial para a análise textual posterior, pois padroniza os dados e facilita a identificação de palavras-chave relevantes.

Além disso, a coluna “*entry-date*” foi tratada para assegurar a uniformidade das datas, convertendo-as para o formato de data padrão (YYYY-MM-DD). Essa padronização é fundamental para permitir uma análise temporal precisa das ideias de negócios. Para a análise, o título e o conteúdo das ideias foram concatenados em uma nova coluna chamada 'texto', proporcionando uma visão integrada de cada ideia. Em seguida, foi realizada a remoção de duplicatas com base nesta coluna 'texto', assegurando a singularidade de cada ideia no conjunto de dados.

Uma etapa adicional de refinamento incluiu a criação de uma lista personalizada de “*stop words*”, contendo termos comuns que poderiam não contribuir significativamente para a distinção entre as ideias de negócios. A exclusão dessas palavras visa focar a análise em termos tematicamente relevantes.

Após essas etapas, obteve-se um conjunto de dados limpo e estruturado, pronto para as fases seguintes de vetorização e análise exploratória. A Figura 4 apresenta o resultado desta etapa de limpeza e preparação dos dados.

Figura 4 – Antes e depois da limpeza e preparação dos dados

```
[15] # Realizando a limpeza dos dados e pré-processamento
data['entry-title'] = data['entry-title'].str.lower().str.replace('[^\w\s]', '', regex=True)
data['entry-content'] = data['entry-content'].str.lower().str.replace('[^\w\s]', '', regex=True)
data['entry-date'] = pd.to_datetime(data['entry-date'], errors='coerce')

# Concatenando título e conteúdo para análise
data['texto'] = data['entry-title'] + ' ' + data['entry-content']

# Removendo duplicatas
data = data.drop_duplicates(subset='texto')

# Visualizando Clusters k-means
data.head()
```

post-categories	entry-title	entry-title href	entry-content	entry-date	texto
0	Technology	Improved Tsunami Warning Systems	With the aid of arms control technology and ar...	April 22, 2023	improved tsunami warning systems with the aid ...
1	Technology	Minimally Invasive Surgical Instruments	One of the challenges of minimally invasive la...	April 17, 2023	minimally invasive surgical instruments one of...
2	Technology	Wearable Plant Patch Monitors	We may be one step closer to using technology ...	April 16, 2023	wearable plant patch monitors we may be one st...
3	Technology	Smart Projector	Though probably best known for its instant cam...	April 10, 2023	smart projector though probably best known for...
4	Technology	M Read Promises Better Land Mine Detection	While metal detectors are useful for detecting ...	April 7, 2023	m read promises better land mine detection whil...

Fonte: elaborado pelo autor

4.2 VETORIZAÇÃO E PRÉ-PROCESSAMENTO

Após a limpeza e preparação dos dados, o conteúdo textual foi submetido a um processo de representação numérica para análise computacional. Para isso, empregou-se a métrica estatística TF-IDF (*Term Frequency-Inverse Document Frequency*), que avalia a relevância de palavras específicas dentro de um documento em relação ao corpus total. Essa métrica ajudou a transformar os

documentos em vetores numéricos num espaço de características, onde cada dimensão corresponde a um termo do corpus, ponderado pelo seu valor TF-IDF.

Na configuração do *TfidfVectorizer*, definiu-se um limite máximo de frequência de documento (*max_df*) de 1.0, que inclui todos os termos independentemente de sua frequência comum, e um limite mínimo (*min_df*) de 2, para excluir palavras raras que poderiam ser consideradas ruídos. Com esses parâmetros, cada ideia de negócio foi transformada em um vetor numérico no espaço de características, refletindo a importância relativa de suas palavras.

Os vetores resultantes, agora numa matriz de documentos normalizados por TF-IDF, facilitaram a análise computacional e estabeleceram a base para os próximos passos analíticos, incluindo o agrupamento das ideias de negócios com o algoritmo *K-Means* e a redução de dimensionalidade com PCA

4.3 REDUÇÃO DE DIMENSIONALIDADE E *CLUSTERING*

Após a vetorização dos dados utilizando TF-IDF, foi realizada a redução de dimensionalidade através da Análise de Componentes Principais (PCA). Este método estatístico possibilitou a transformação do espaço de alta dimensionalidade dos vetores TF-IDF em um novo espaço de dimensões reduzidas. A escolha de dois componentes principais foi motivada pela necessidade de uma visualização bidimensional, que auxilia na interpretação e identificação de padrões entre as ideias de negócios.

A análise preliminar das componentes principais revelou agrupamentos naturais, sugerindo que ideias de negócios com conteúdos textuais semelhantes tendem a se agrupar em espaço reduzido. Em seguida, aplicou-se o algoritmo K-Means diretamente aos dados transformados pela PCA, segmentando as ideias de negócios em cinco *clusters* distintos. A escolha de cinco *clusters* foi determinada como um equilíbrio entre representar adequadamente a diversidade dos dados e manter a clareza visual da análise. Essa quantidade foi considerada suficiente para abranger a variedade temática sem criar uma superposição excessiva ou poluição visual, que poderia ocorrer com um número maior de *clusters*. Cabe ressaltar que essa escolha não foi baseada em uma métrica específica, mas sim na necessidade de uma representação clara e gerenciável dos dados no espaço bidimensional.

Adicionalmente, para uma análise mais detalhada, o K-Means foi também aplicado à matriz TF-IDF original, permitindo a extração de palavras-chave significativas de cada *cluster*. Esta análise ofereceu uma compreensão abrangente tanto da distribuição espacial quanto dos temas predominantes em cada *cluster*, evidenciado na Figura 5.

Figura 5 – Visualização dos clusters após execução do K-Means na matriz TF-IDF

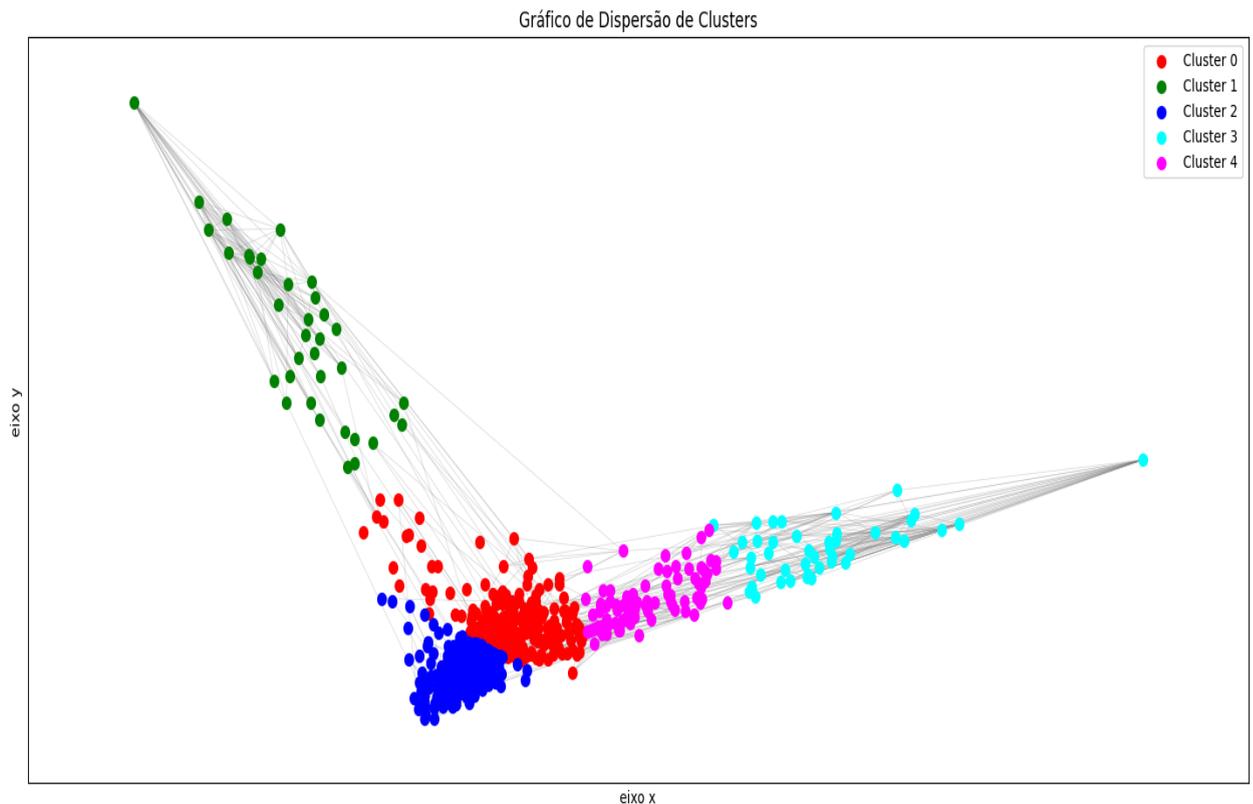
Cluster	Keywords
0	camera, time, tech, drone, business, world, smart, company, robot, technology
1	world, desktop, printers, resin, print, objects, printed, printer, printing, 3d
2	sensor, intelligence, artificial, team, researchers, technology, people, scientists, developed, university
3	university, charging, glass, electricity, lithium, power, batteries, battery, energy, solar
4	filter, solar, electricity, sunlight, material, liquid, air, cooling, heat, water

Fonte: elaborado pelo autor

Os resultados do agrupamento, visualizados na Figura 6 como um gráfico de dispersão, oferecem uma perspectiva reveladora sobre a distribuição e as relações das ideias de negócios no domínio da tecnologia.

A Figura 6 demonstra a distribuição das ideias de negócios mapeadas nas duas principais componentes resultantes da análise PCA. A representação colorida destaca a existência de cinco *clusters* distintos, identificados com cores específicas que facilitam a interpretação visual: vermelho para o *cluster* 0, verde para o *cluster* 1, azul para o *cluster* 2, ciano para o *cluster* 3 e magenta para o *cluster* 4.

Figura 6 – Gráfico de dispersão de clusters



Fonte: elaborado pelo autor

O *cluster 0*, em vermelho, se destaca como o mais numeroso, abrigando 465 ideias. A densidade de pontos neste *cluster* sugere uma temática amplamente compartilhada, refletindo possivelmente uma convergência em torno de temas como inteligência artificial ou automação, dada a sua predominância no setor tecnológico. O *cluster 1*, colorido em verde e contendo 45 ideias, embora seja o menor em termos de número, apresenta uma distinção notável no espaço de plotagem. Isso pode indicar um nicho específico de inovação ou uma abordagem diferenciada dentro do campo da tecnologia, talvez focando em áreas emergentes ou menos exploradas. O *cluster 2* (azul) e o *cluster 3* (ciano), com 238 e 112 ideias respectivamente, ocupam um espaço intermediário em termos de tamanho. A distribuição espacial desses *clusters* pode apontar para uma variedade de subtemas ou para uma mistura de conceitos inovadores que ainda mantêm certa coesão temática. O *cluster 4*, representado em magenta e composto por 61 ideias, reflete uma diversidade interna mais acentuada. Este *cluster* pode estar abrangendo ideias de negócios que desafiam as convenções ou que combinam múltiplos campos tecnológicos de maneira única.

A análise espacial proporcionada pelo gráfico de dispersão não apenas ilustra a separação clara dos *clusters*, mas também revela as áreas de sobreposição. Estas áreas de intersecção são particularmente interessantes, pois podem indicar a convergência de diferentes campos tecnológicos ou a evolução de novas tendências que abrangem múltiplos domínios.

4.4 ANÁLISE EXPLORATÓRIA

Nesta etapa da pesquisa, foi realizada a análise exploratória dos *clusters* formados para extrair *insights* sobre as tendências e características das ideias de negócios na área da tecnologia.

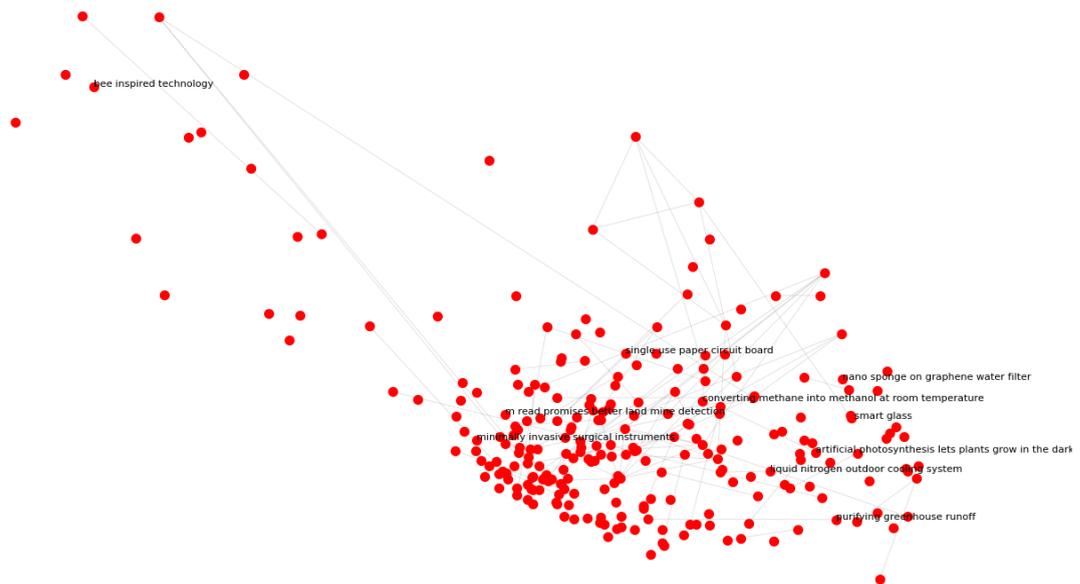
4.4.1 Análise Detalhada dos *clusters* de Ideias de Negócios

Além do gráfico de dispersão de todos os *clusters*, uma visualização complementar de cada um deles, separadamente, oferece uma análise mais profunda sobre as tendências e temáticas predominantes em cada agrupamento.

Na Figura 7, que representa o *cluster* 0, observa-se uma concentração de ideias relacionadas a inovações médicas e soluções sustentáveis. Títulos como "*Minimally Invasive Surgical Instruments*" (Instrumentos Cirúrgicos Minimamente Invasivos) e "*Smart Glass*" (Vidro Inteligente) indicam um foco em tecnologias que buscam melhorar procedimentos médicos e incrementar a eficiência energética. Outras ideias como "*Bee Inspired Technology*" (Tecnologia Inspirada em Abelhas) e "*Nano Sponge on Graphene Water Filter*" (Esponja Nano em Filtro de Água de Grafeno) refletem um interesse em soluções sustentáveis e ambientalmente amigáveis.

Figura 7 – Rede de ideias do *cluster 0*

Rede de Clusters de Ideias - Cluster 0

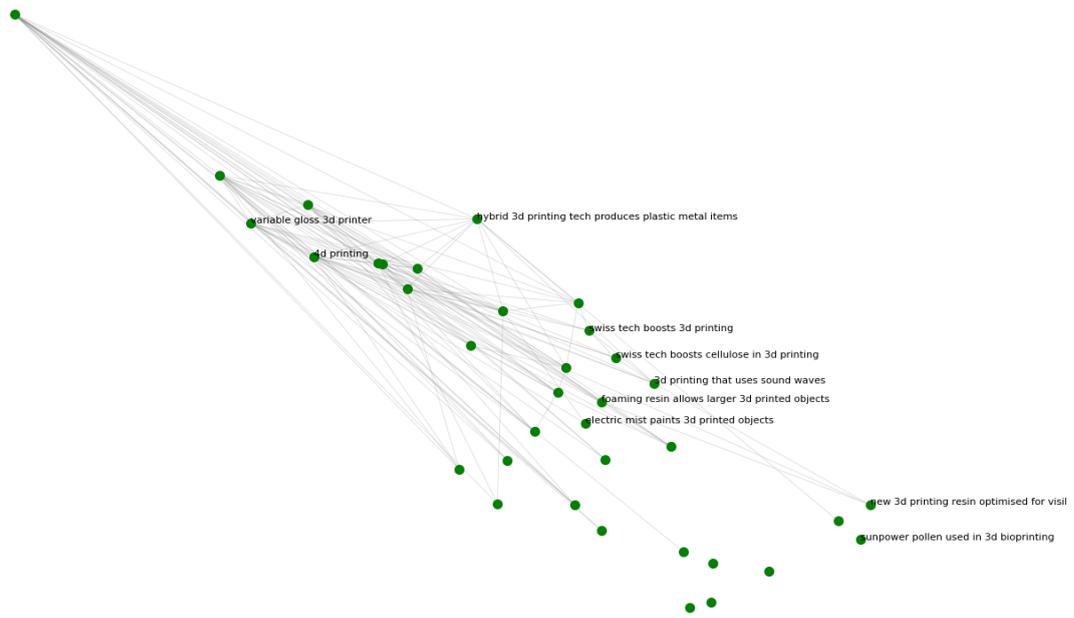


Fonte: elaborado pelo autor

Ilustrado na Figura 8, o *cluster 1* destaca-se pelo seu foco na impressão 3D e *bioprinting*. Ideias como "3D Printing that Uses Sound Waves" (Impressão 3D que Usa Ondas Sonoras) e "4D Printing" (Impressão 4D) mostram inovações nesta tecnologia. O uso de novos materiais, como demonstrado em "Sunpower Pollen Used in 3D Bioprinting" (Pólen Sunpower Usado em Bioimpressão 3D), e a otimização de processos, como em "New 3D Printing Resin Optimised for Visible Light" (Nova Resina de Impressão 3D Otimizada para Luz Visível), ilustram uma evolução significativa na área.

Figura 8 – Rede de ideias do *cluster 1*

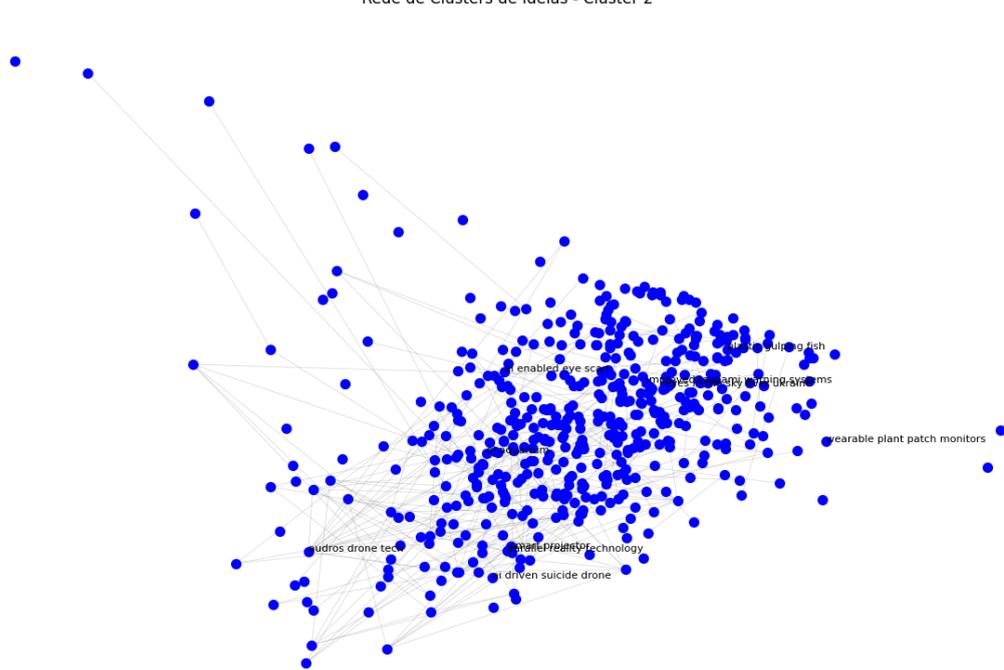
Rede de Clusters de Ideias - Cluster 1



Fonte: elaborado pelo autor

Na Figura 9, que ilustra o *cluster 2*, destacam-se as ideias que combinam tecnologias ambientais com aplicações de inteligência artificial. Títulos como "*Improved Tsunami Warning Systems*" (Sistemas Aperfeiçoados de Alerta de Tsunami) e "*AI Aquarium*" (Aquário com IA) destacam o uso de tecnologia avançada em aplicações práticas. Outros títulos, como "*Plastic Gulping Fish*" (Peixe que Engole Plástico) e "*AI Enabled Eye Scan*" (Varredura Ocular Habilitada por IA), sugerem um foco em soluções tecnológicas para problemas ambientais e de saúde.

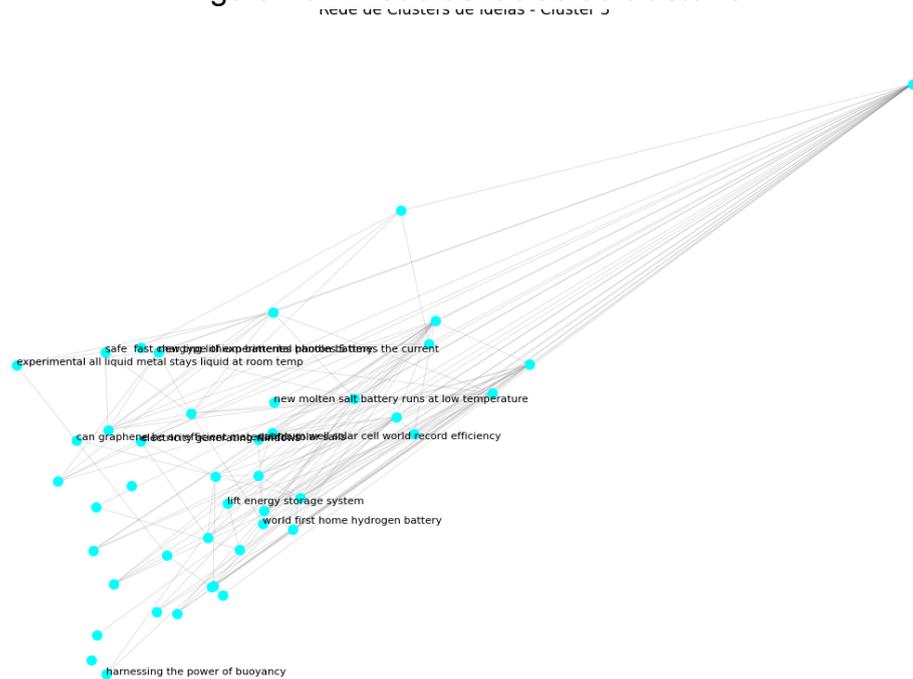
Figura 9 – Rede de ideias do *cluster 2*



Fonte: elaborado pelo autor

O *cluster 3*, apresentado na Figura 10, concentra-se em energia renovável e soluções de armazenamento de energia. Títulos como "*Electricity Generating Windows*" (Janelas Geradoras de Eletricidade) e "*Quantum Well Solar Cell World Record Efficiency*" (Eficiência Recorde Mundial em Célula Solar de Poço Quântico) indicam um forte interesse em fontes de energia sustentáveis. Além disso, ideias como "*New Molten Salt Battery Runs at Low Temperature*" (Nova Bateria de Sal Fundido Funciona em Baixas Temperaturas) e "*World First Home Hydrogen Battery*" (Primeira Bateria Doméstica de Hidrogênio do Mundo) mostram avanços no armazenamento de energia.

Figura 10 – Rede de ideias do *cluster 3*

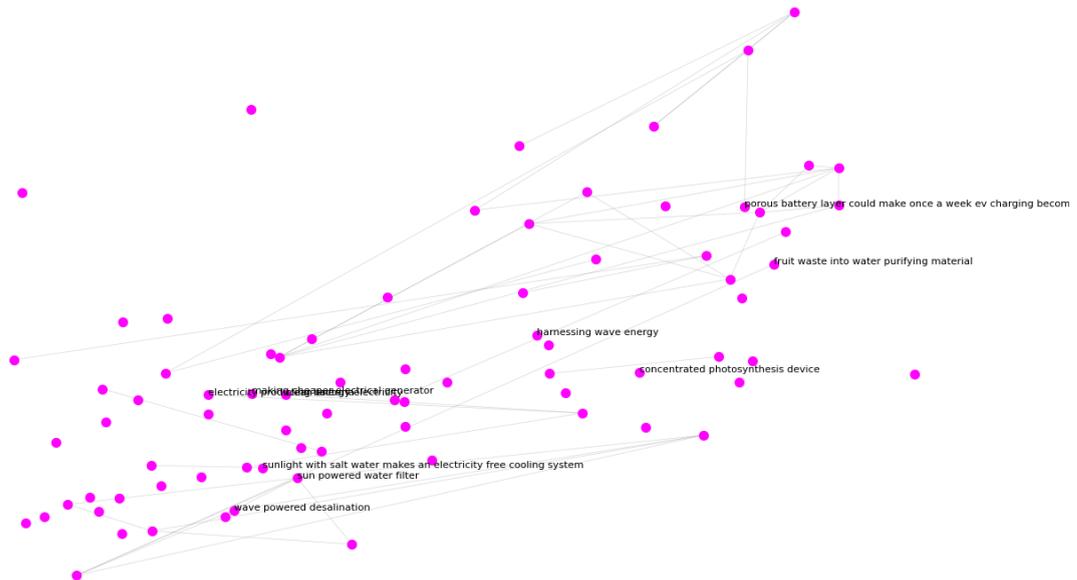


Fonte: elaborado pelo autor

Finalmente, o *cluster 4*, ilustrado na Figura 11, abrange ideias relacionadas à purificação de água e geração de energia. Títulos como "*Fruit Waste into Water Purifying Material*" (Transformação de Resíduos de Frutas em Material Purificador de Água) e "*Wave Powered Desalination*" (Dessalinização Movida a Ondas) mostram uma abordagem criativa no uso de recursos naturais. Ideias como "*Electricity Producing Bacteria*" (Bactérias Produtoras de Eletricidade) e "*Sunlight with Salt Water Makes an Electricity-Free Cooling System*" (Luz Solar com Água Salgada Cria um Sistema de Resfriamento Sem Eletricidade) sugerem inovações que combinam tecnologia e sustentabilidade.

Figura 11 – Rede de ideias do *cluster 4*

Rede de Clusters de Ideias - Cluster 4



Fonte: elaborado pelo autor

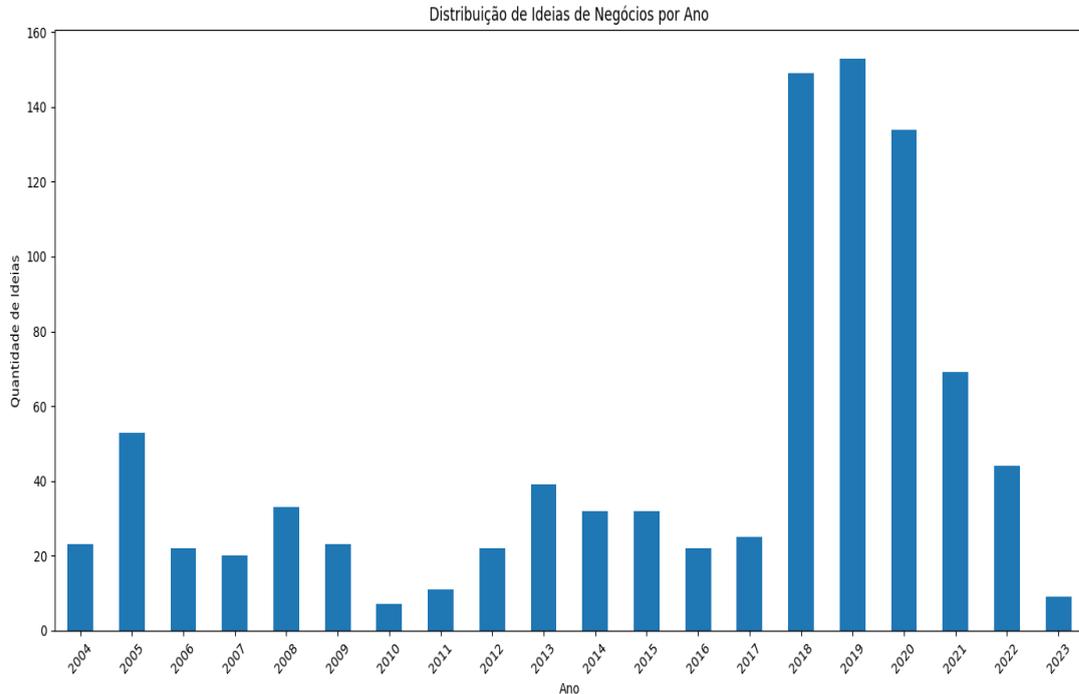
Cada *cluster* representa uma área distinta de foco e inovação, desde avanços médicos e sustentabilidade até impressão 3D e energia renovável. Esta variedade demonstra a amplitude e a diversidade de ideias exploradas no campo tecnológico.

4.4.2 Análise Temporal das Ideias de Negócios e Confirmação de Ideias Antigas

Na análise exploratória das ideias de negócios ao longo dos anos, observou-se uma distribuição temporal que reflete tendências e mudanças no cenário de inovação tecnológica. Uma elevação significativa no volume de ideias foi registrada a partir de 2018, indicando um possível aumento no interesse por inovações ou uma alteração na metodologia de coleta do site “coolbusinessIdeas.com”. Embora haja oscilações nos anos anteriores, é a partir de 2017 que se evidencia uma tendência de crescimento mais acentuada. O ano de 2020, em particular, mostrou um pico no número de ideias, provavelmente em resposta aos desafios impostos pela pandemia global, que estimulou uma onda de inovação e a busca por novos modelos de negócios adaptativos. Esta tendência

pode ser visualizada na Figura 12, que apresenta um gráfico de barras das ideias de negócios por ano.

Figura 12 – Gráfico de barra de Ideias de negócios por Ano



Fonte: elaborado pelo autor

Complementando a análise temporal, investigou-se a relevância contínua de ideias de negócios concebidas entre 2004 e 2010, com o objetivo de avaliar seu sucesso nos dias atuais. Esta análise não apenas corrobora o estudo de *insights* para a geração de novas ideias, mas também destaca a importância de inovações passadas no cenário tecnológico contemporâneo. Estas cinco ideias exemplificam como os *insights* passados podem informar e influenciar o desenvolvimento de negócios tecnológicos atuais:

- **IVF Freezing (2006-05-09):** A tecnologia de congelamento para fertilização in vitro (IVF) tem se mantido como uma prática padrão na medicina reprodutiva, representando um sucesso duradouro. A conservação de material genético tem permitido avanços significativos no tratamento da infertilidade, provando ser uma inovação com impacto profundo e de longa duração.
- **Business Continuity Challenges in the IT Field (2005-06-08):** A resiliência e continuidade dos negócios no setor de TI, destacadas pelo

Gartner, tornaram-se imperativos na era da transformação digital. Com o aumento do trabalho remoto e eventos globais disruptivos, as estratégias de continuidade de negócios são agora mais críticas do que nunca, demonstrando a presciência e relevância desta ideia inicial.

- ***Waterproofing the World: P2i Plasma Technology (2008-09-21)***: A demanda por impermeabilização estendeu-se além do vestuário e dispositivos eletrônicos, abrangendo uma variedade de setores. Esta inovação tecnológica previu a necessidade de produtos duráveis e resistentes às intempéries, o que é atualmente uma expectativa de mercado.
- ***Speak your Mind (2008-04-05)***: As interfaces cérebro-computador que facilitam a comunicação para pessoas com deficiências continuam a evoluir. Este campo de pesquisa está crescendo exponencialmente, com aplicações potenciais que vão desde a assistência médica até a integração com tecnologias de consumo, confirmando a relevância e o potencial de longo prazo desta ideia.
- ***Learning Ecosystem Through Tech (2005-04-23)***: A integração da tecnologia na educação transformou-se em uma prática pedagógica fundamental. Plataformas de e-learning, realidade aumentada e virtual são agora ferramentas essenciais no ensino, tanto em ambientes formais quanto informais de aprendizado. Esta ideia, que uma vez foi considerada pioneira, é agora uma realidade no cenário educacional.

Esta análise demonstra a utilidade de examinar ideias anteriores como uma ferramenta para prever e entender tendências futuras, e oferece uma confirmação empírica para a abordagem metodológica adotada nesta pesquisa.

4.4.3 Análise de Palavras-chave e Tendências Temáticas

A investigação das palavras-chave mais representativas dos *clusters* identificados revelou temas predominantes nas ideias de negócios:

No *cluster* 0, palavras como "*camera*", "*drone*" e "*robot*" destacam uma forte inclinação para a inovação em tecnologias de vigilância e automação. conforme ilustrado na Figura 13. Este *cluster* sugere uma integração crescente de dispositivos

inteligentes na vida cotidiana, refletindo um mercado em evolução que busca soluções avançadas em robótica e tecnologia de ponta.

Figura 13 – Nuvem de palavras referente ao cluster 0



Fonte: elaborado pelo autor

O *cluster 1* é caracterizado por termos como "*desktop*", "*printers*" e "*3d*", apontando para um desenvolvimento significativo no campo da impressão 3D, como evidenciado na Figura 14. A ênfase em "*resin*" e "*printing*" indica inovações em materiais e métodos de fabricação, transformando a maneira como objetos são concebidos e produzidos.

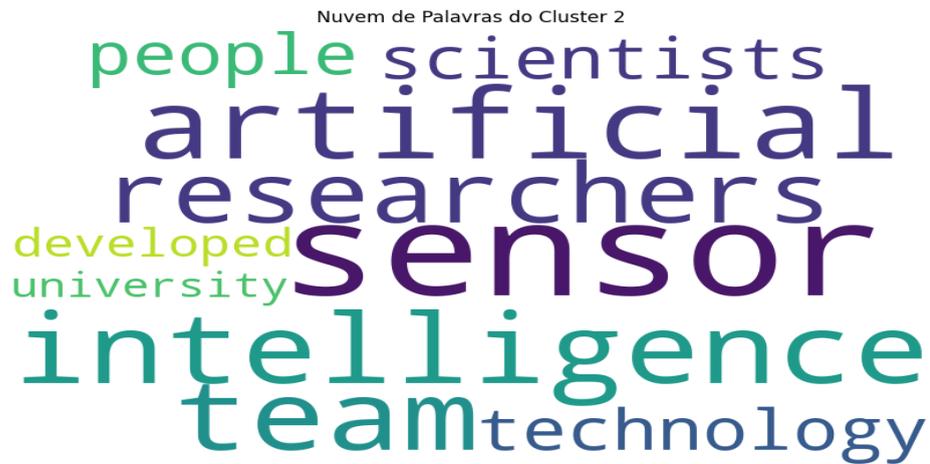
Figura 14 – Nuvem de palavras referente ao cluster 1



Fonte: elaborado pelo autor

O *cluster 2*, apresentado na Figura 15, destaca-se com palavras como "sensor", "artificial intelligence" e "university", evidenciando um compromisso com a pesquisa e o desenvolvimento, especialmente em projetos que combinam tecnologia social e ambiental com inteligência artificial. Este *cluster* também sublinha a importância das colaborações entre universidades e pesquisadores na inovação tecnológica.

Figura 15 – Nuvem de palavras referente ao cluster 2



Fonte: elaborado pelo autor

No *cluster 3*, termos como "solar", "energy" e "batteries", apresentados na Figura 16, indicam um foco nas energias renováveis e em tecnologias de armazenamento de energia. Esta tendência para soluções sustentáveis reflete a crescente preocupação global com eficiência energética e a redução do impacto ambiental.

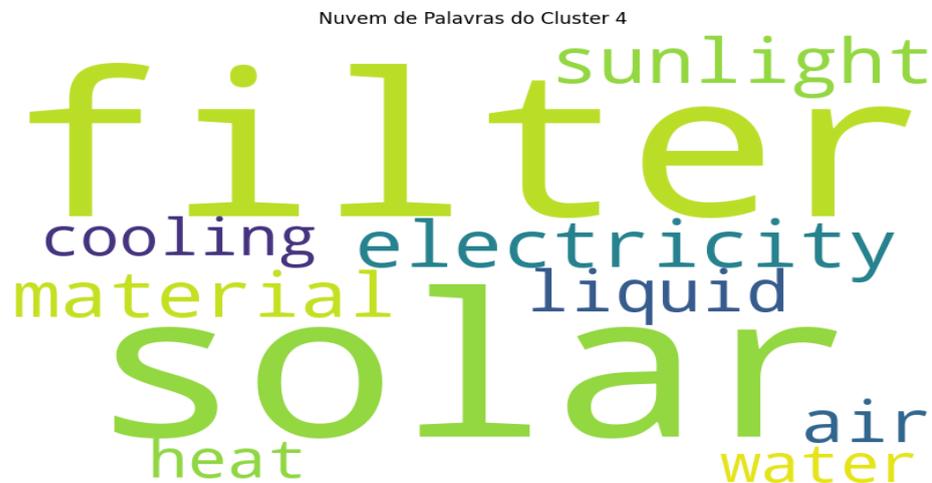
Figura 16 – Nuvem de palavras referente ao cluster 3



Fonte: elaborado pelo autor

Por fim, o *cluster* 4, representado na Figura 17, enfatiza inovações ambientais, com palavras-chave como “*electricity*”, “*filter*”, “*water*” e “*cooling*”. Esses termos apontam para avanços significativos em tecnologias ambientais e eficiência energética. A presença de “*electricity*” e “*filter*” sugere um foco em soluções para purificação de água e gestão eficiente de recursos energéticos, enquanto “*cooling*” indica desenvolvimentos em sistemas de climatização mais sustentáveis e eficientes.

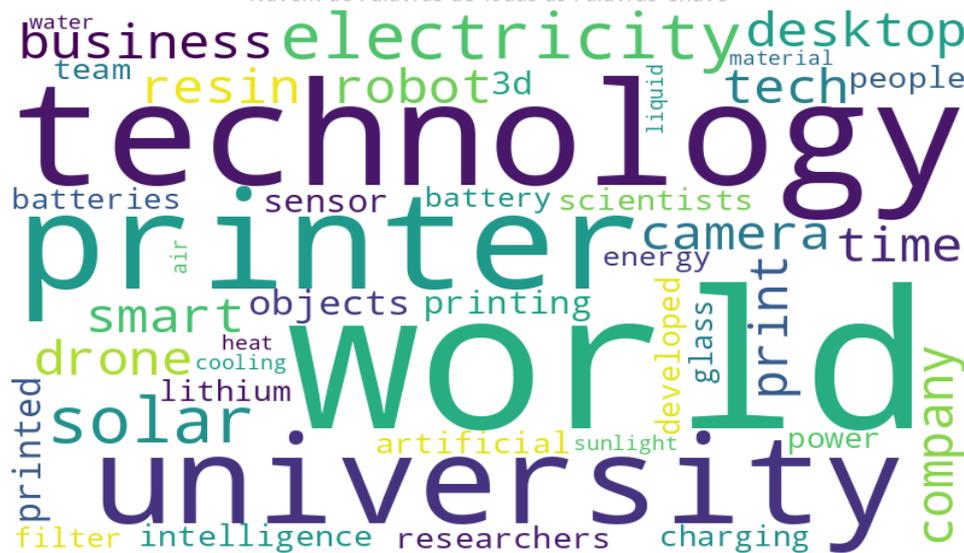
Figura 17 – Nuvem de palavras referente ao cluster 4



Fonte: elaborado pelo autor

A investigação das palavras-chave mais representativas dos clusters identificados revelou temas predominantes nas ideias de negócios. Para capturar a essência das tendências em todos os *clusters*, uma nuvem de palavras combinada foi gerada, ilustrada na Figura 18, destacando os termos mais frequentes em todo o conjunto de ideias dos cinco *clusters*.

Figura 18 – Nuvem de palavras unindo os cinco clusters



Fonte: elaborado pelo autor

Em suma, a análise dos *clusters* fornece *insights* significativos sobre as direções predominantes das inovações tecnológicas. As tendências identificadas estão alinhadas com um mercado em constante evolução e com uma sociedade cada vez mais consciente dos desafios ambientais e da necessidade de desenvolvimento sustentável.

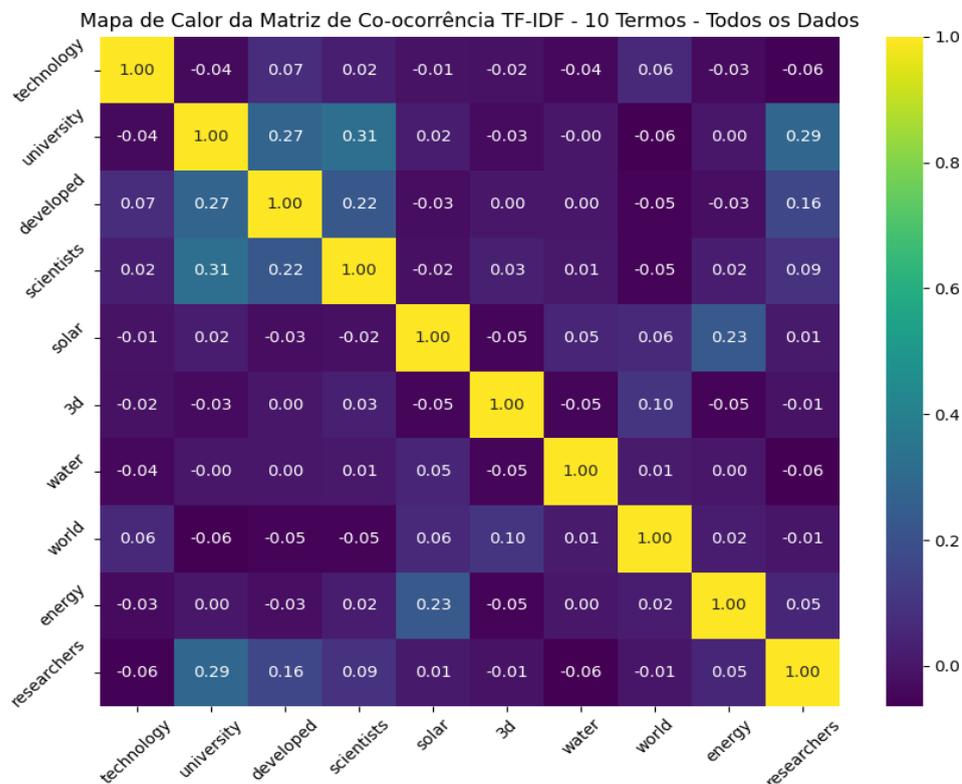
4.4.4 Análise de Palavras-chave e Mapas de Calor

Nesta etapa do estudo, foram exploradas as conexões temáticas entre as ideias de negócios tecnológicos através da técnica de mapas de calor, fundamentada na correlação de palavras-chave extraídas da matriz TF-IDF. Esta abordagem visual foi aplicada em três conjuntos de dados distintos: um abrangendo o período completo de análise (Figura 19), o período da pandemia (2019-2020, Figura 20), e os últimos três anos (2021-2023, Figura 21).

A análise dos mapas de calor, baseada na correlação das palavras-chave durante a pandemia (2019-2020) e nos últimos três anos (2021-2023), revela padrões intrigantes e coerentes com as tendências previamente observadas. Durante a pandemia, as correlações mais fortes entre palavras como "researchers", "university", "developed", e "scientists" destacam uma concentração significativa em pesquisa acadêmica e desenvolvimento científico. Esta tendência reflete uma

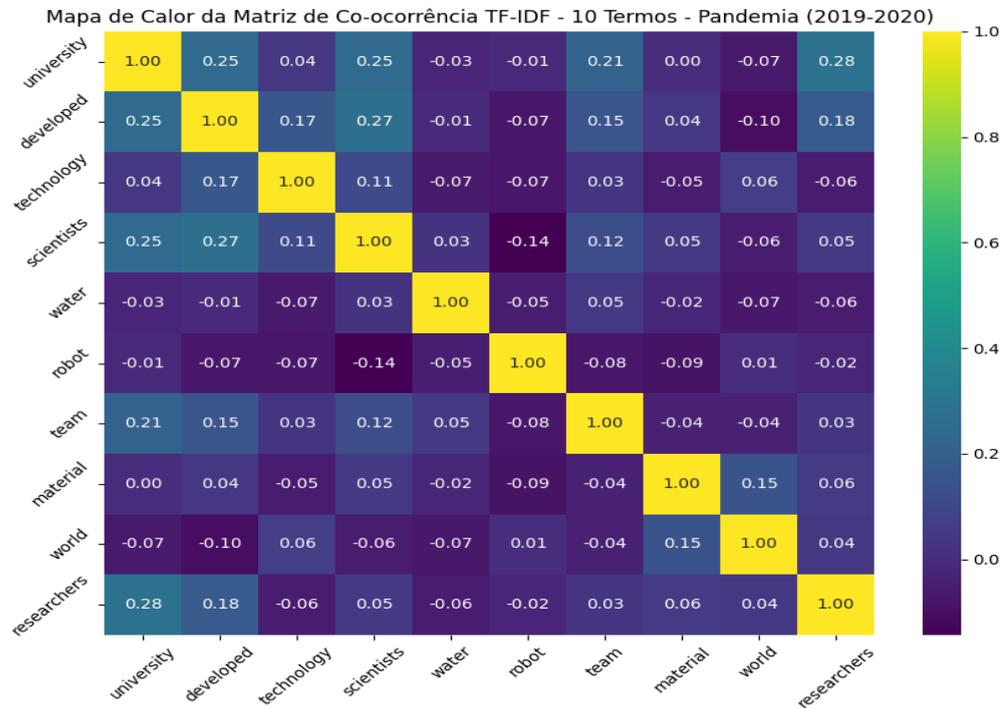
resposta direta aos desafios globais impostos pela pandemia, com um foco acentuado na colaboração entre instituições acadêmicas e cientistas para inovações rápidas e eficazes. Nos últimos três anos, essa tendência de colaboração intensiva e foco em pesquisa e desenvolvimento permaneceu predominante, como evidenciado pela continuidade das fortes correlações entre termos semelhantes. Este padrão sublinha o papel crucial das universidades e cientistas no avanço contínuo da tecnologia e inovação. Além disso, a correlação entre "technology" e "team" sugere um movimento em direção a uma abordagem mais integrada e multidisciplinar no desenvolvimento tecnológico. Esses *insights* reforçam a ideia de que tanto durante quanto após a pandemia, houve um impulso significativo na colaboração e inovação em pesquisa. As tendências indicam um mercado cada vez mais focado em soluções inovadoras que não apenas respondem a desafios imediatos, mas também incorporam uma visão de longo prazo para sustentabilidade e desenvolvimento tecnológico.

Figura 19 – Mapa de calor da correlação de termos normalizados por TF-IDF - 10 termos - todos os dados



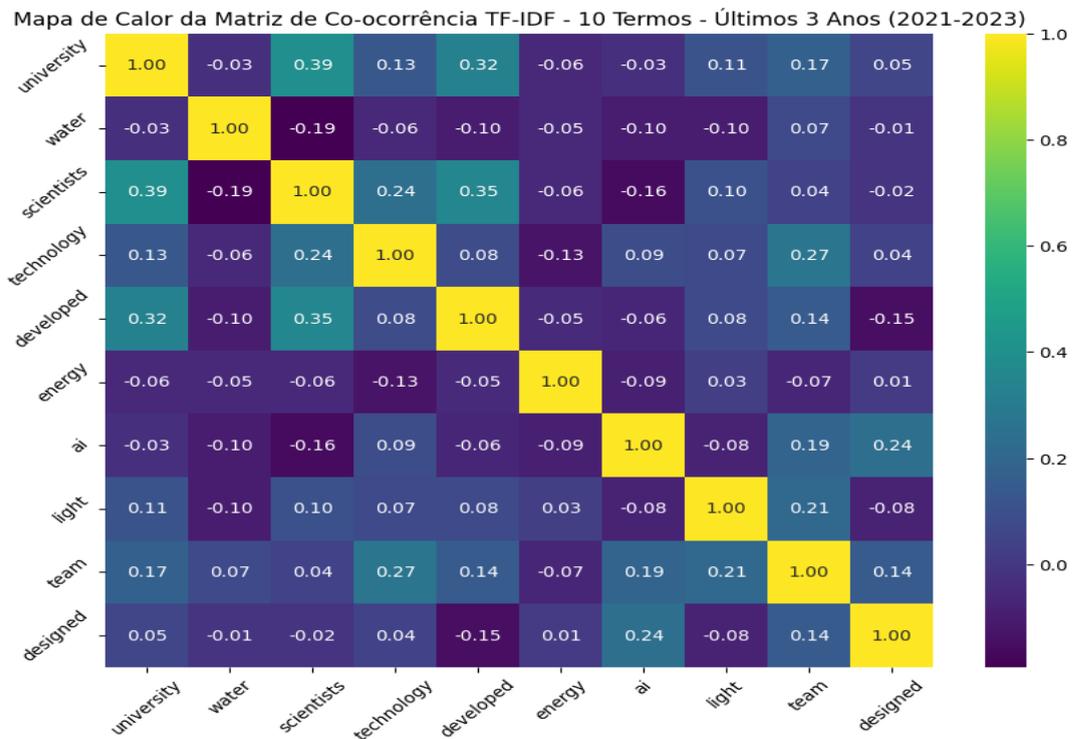
Fonte: elaborado pelo autor

Figura 20 – Mapa de calor da correlação de termos normalizados por TF-IDF - 10 termos - pandemia (2019-2020)



Fonte: elaborado pelo autor

Figura 21 – Mapa de calor da correlação de termos normalizados por TF-IDF - 10 termos - últimos 3 anos (2021-2023)



Fonte: elaborado pelo autor

A análise sugere ainda que ocorre certa integração entre inteligência artificial e sustentabilidade. A constante correlação entre estas áreas aponta para uma sinergia crescente entre tecnologia avançada e práticas sustentáveis, oferecendo caminhos promissores para o futuro da inovação tecnológica.

Essa tendência é acompanhada pelo destaque contínuo em tecnologias autônomas, sugerindo que o mercado está cada vez mais orientado para soluções que combinam autonomia, eficiência e responsabilidade ambiental. Assim, a análise dos mapas de calor reforça a noção de que se está diante de um cenário tecnológico dinâmico e em evolução, onde a inovação consciente e colaborativa é fundamental para enfrentar os desafios atuais e futuros.

4.4.5 Insights Extraídos das Visualizações

A análise aprofundada das ideias de negócios tecnológicos, utilizando dados do site "CoolBusinessIdeas.com", revelou uma teia complexa de interconexões temáticas. Através de visualizações como mapas de calor, gráficos de redes e nuvens de palavras, descobriu-se um panorama intrincado de tendências e padrões emergentes, que são fundamentais para compreender a evolução da inovação tecnológica. Um tema recorrente nas visualizações foi a fusão entre inteligência artificial (IA) e sustentabilidade. Esta tendência reflete uma direção inovadora onde a IA não é apenas aplicada para avanços tecnológicos, mas também integrada a práticas sustentáveis. Os mapas de calor mostraram uma correlação persistente entre termos como "*autonomous*", "*technology*", "*battery*", "*electric*" e "*renewable*", sugerindo uma abordagem holística que combina automação e eficiência energética com soluções de energia renovável.

Nos gráficos de redes, observou-se uma interconexão significativa entre termos relacionados à autonomia, como "*autonomous*" e "*technology*", e soluções sustentáveis, como "*solar*" e "*energy*". Isso indica que a inovação em autonomia está sendo cada vez mais aplicada em contextos diversificados, incluindo soluções de energia renovável. A frequente correlação de termos como "*researchers*", "*university*", "*developed*" e "*scientists*" aponta para um aumento na colaboração acadêmica e uma maior ênfase na pesquisa científica aplicada como motor de inovação.

As nuvens de palavras de cada *cluster* revelaram focos temáticos distintos, mas com sobreposições reveladoras. A recorrência de termos como "battery", "electric" e "renewable" em vários *clusters* sublinha um movimento em direção a soluções integradas, onde armazenamento de energia e energias renováveis são vistos como componentes complementares em estratégias energéticas mais abrangentes.

Esta análise indica que a inovação tecnológica está evoluindo para uma integração mais sistemática e holística de diferentes campos. A convergência entre IA, tecnologias autônomas e práticas sustentáveis sinaliza um futuro onde a inovação será cada vez mais avaliada pela sua capacidade de fornecer soluções holísticas e sustentáveis. A sinergia entre estes campos oferece um caminho promissor para o desenvolvimento de estratégias de negócios que equilibram avanços tecnológicos com responsabilidades ambientais e sociais.

4.4.6 Implicações para a Inovação e Estratégia de Negócios

As análises dos dados de negócios tecnológicos trazem à luz importantes implicações para as estratégias de inovação e negócios no atual contexto de mercado. A interseção crescente entre Inteligência Artificial e sustentabilidade, assim como a persistência das tecnologias autônomas e de armazenamento de energia, apontam para direções estratégicas cruciais que as organizações podem seguir para se manterem competitivas e relevantes.

Diante desse cenário, as organizações devem considerar as seguintes abordagens:

- **Integração de IA com Sustentabilidade:** O desenvolvimento de produtos e soluções que aliam avanços técnicos em IA com compromissos ambientais é essencial. Investir em pesquisa e desenvolvimento que enfoca a aplicação da IA para melhorar a sustentabilidade é uma estratégia vital. Isso não apenas amplia o espectro de inovação tecnológica, mas também assegura que tais inovações contribuem de maneira significativa para os desafios ambientais contemporâneos.
- **Parcerias Interdisciplinares e Inter-Setoriais:** Fomentar colaborações que transcendam fronteiras disciplinares e setoriais é crucial. A exploração da

sinergia entre diferentes áreas de especialização permite abordar problemas complexos de forma mais holística e inovadora. Parcerias entre indústrias, universidades e institutos de pesquisa podem ser particularmente frutíferas, potencializando a pesquisa aplicada e a inovação.

- **Revisão e Adaptação de Modelos de Negócios:** As organizações precisam constantemente revisar e adaptar seus modelos de negócios para incorporar práticas sustentáveis. Isso implica alinhar inovações com valores de responsabilidade corporativa e as demandas crescentes do mercado por soluções ambientalmente conscientes. Uma abordagem de negócios que equilibre inovação e sustentabilidade pode não apenas aumentar a competitividade, mas também promover uma imagem corporativa positiva.
- **Vigilância Estratégica Ativa:** Manter-se atento às tendências emergentes e adaptar-se proativamente às mudanças no cenário tecnológico é fundamental. Utilizar análises de dados avançadas para identificar novas oportunidades e prever tendências futuras permite que as empresas se antecipem a mudanças de mercado e se posicionem estrategicamente. Isso pode incluir a exploração de novos mercados, a diversificação de ofertas de produtos ou serviços, ou a adaptação a novos regulamentos e expectativas dos consumidores.

Em resumo, os *insights* obtidos a partir das análises de dados destacam a necessidade de uma abordagem de negócios que seja inovadora, sustentável e adaptável. As organizações que conseguirem integrar inteligência artificial com práticas sustentáveis, fomentar colaborações interdisciplinares, adaptar seus modelos de negócios e manter uma vigilância estratégica ativa estarão melhor posicionadas para prosperar em um ambiente de negócios em constante evolução.

4.4.7 Limitações do Estudo e Considerações Futuras

Apesar dos *insights* valiosos, este estudo não está isento de limitações. A base de dados do *site* “CoolBusinessIdeas.com” pode não capturar a totalidade do ecossistema de inovações tecnológicas, podendo haver viés na seleção das ideias publicadas. Além disso, a análise foi limitada às informações disponíveis na

descrição das ideias de negócios, que podem não fornecer um quadro completo do potencial de mercado ou da viabilidade técnica das inovações propostas.

Para pesquisas futuras, seria proveitoso expandir o conjunto de dados para incluir múltiplas fontes e abranger uma variedade mais ampla de inovações tecnológicas. Além disso, seria valioso aplicar outras técnicas de análise, a fim de extrair *insights* ainda mais profundos e proporcionar uma compreensão mais dinâmica das trajetórias de inovação.

Concluindo, enquanto este estudo fornece uma base para entender as tendências atuais e emergentes em tecnologia, a constante evolução do mercado exige que tais análises sejam realizadas regularmente para manter a relevância e precisão dos *insights* para estratégias de negócios e inovação.

5 CONSIDERAÇÕES FINAIS

Este trabalho apresentou uma análise detalhada de ideias de negócios da área de tecnologia, coletadas do site CoolBusinessIdeas.com, com o intuito de discernir padrões e tendências que informem estratégias de inovação e gestão. A pesquisa percorreu um caminho analítico, desde a limpeza e preparação dos dados até a aplicação de técnicas como TF-IDF, K-Means e PCA, para vetorização, agrupamento e redução de dimensionalidade dos dados. A análise temporal ressaltou na identificação de um aumento acentuado no número de ideias inovadoras a partir de 2018, com um pico em 2020, provavelmente em resposta aos desafios da pandemia.

A relevância contínua de conceitos de negócios surgidos nas primeiras etapas do período estudado foi confirmada, demonstrando que certas inovações mantêm sua importância e aplicabilidade ao longo do tempo. Isso sublinha a importância de visitar e reavaliar ideias anteriores, pois elas podem servir como fundamento ou inspiração para novas inovações.

A pesquisa atendeu ao seu objetivo geral de analisar as tendências em ideias de negócios tecnológicos, fornecendo *insights* claros e quantificáveis para direcionar a inovação e a estratégia de negócios. Os objetivos específicos foram igualmente atendidos: a identificação de padrões temporais, a confirmação de ideias de negócios antigos e a extração de temas predominantes entre os *clusters*. A análise de palavras-chave e tendências temáticas revelou uma forte inclinação para automação, inteligência artificial, sustentabilidade e inovações energéticas.

Este estudo também trouxe contribuições para o campo da análise de dados e visualização no contexto de negócios tecnológicos. Demonstrou-se que as técnicas de processamento de dados e aprendizado de máquina podem ser aplicadas com sucesso para extrair *insights* valiosos de conjuntos de dados textuais extensos. Logo, as visualizações criadas a partir da análise dos dados não apenas ilustraram as tendências de forma clara e intuitiva, mas também facilitaram a compreensão de conceitos complexos.

Esta pesquisa pode contribuir com as empresas e empreendedores que buscam navegar no cenário tecnológico em constante evolução. As descobertas deste estudo oferecem uma base para decisões estratégicas e desenvolvimento de políticas relacionadas à inovação tecnológica. Ao mesmo tempo, a metodologia

adotada pode ser replicada ou adaptada para futuras pesquisas em outras áreas do conhecimento ou setores de negócios.

As limitações identificadas destacam oportunidades para pesquisas futuras. A expansão do conjunto de dados e a incorporação de técnicas analíticas adicionais poderiam enriquecer ainda mais o entendimento das tendências em inovações tecnológicas. A necessidade de atualização contínua das análises reflete a natureza dinâmica do campo tecnológico, enfatizando a importância de se manter constantemente atento às mudanças do mercado para sustentar a relevância e aplicabilidade das descobertas para inovação e estratégia de negócios.

Para pesquisas futuras, identificou-se o potencial em explorar *Large Language Models* (LLMs) e representações vetoriais densas (*embeddings*). Estas técnicas permitem uma representação semântica mais rica de unidades de informação, que podem ser termos, sentenças, textos, imagens, vídeos, ou mesmo subgrafos. Especialmente em um contexto de inovação tecnológica, *embeddings* oferecem uma maneira avançada de capturar as nuances e a complexidade das ideias de negócios. Além disso, os grafos de conhecimento emergem como uma ferramenta promissora para suportar áreas como previsão tecnológica (*technology forecasting*), permitindo a visualização e a análise de conexões entre diferentes conceitos e inovações. Esta abordagem pode ser particularmente útil para empresas e empreendedores que buscam se manter à frente no cenário tecnológico dinâmico. Incorporando essas técnicas avançadas em estudos futuros, será possível aprofundar o entendimento das tendências atuais e antecipar a direção de futuras inovações.

REFERÊNCIAS

- ADAMS, R., BESSANT, J., PHELPS, R. **Innovation management measurement: a review**. International Journal of Management Reviews, vol. 8, nº.1, p. 21–47, 2006.
- ALHADAD, Sakinah SJ. **Visualizing data to support judgment, inference, and decision making in learning analytics**: Insights from cognitive psychology and visualization science. Journal of Learning Analytics, v. 5, n. 2, p. 60–85-60–85, 2018.
- ANDRIENKO, G., ANDRIENKO, N. **Interactive Cluster analysis of diverse types of spatiotemporal data**. ACM Sigkdd Explorations Newsletter, New York, v. 11, nº 2, p. 19-28, 2009.
- BAKKER, Hendrik Jan. **Idea management**: Unraveling creative processes in three professional organizations. 2010.
- BAREGHEH, A.; ROWLEY, J.; SAMBROOK, S. **Towards a multidisciplinary definition of innovation**. Management Decision, United Kingdom, v. 47, n .8, p.1323-1339, 2009.
- BJÖRK, J.; BOCCARDELLI, P.; MAGNUSSON, M. G. **Ideation capabilities for continuous innovation**. Creativity & Innovation Management, Malden, v. 19, n. 4, p. 385-396, 2010.
- BORGES, Livia de Oliveira. **Significados do trabalho e do dinheiro**: uma perspectiva histórica e cultural. Psicologia: Ciência e Profissão, v. 35, n. 1, p. 66-81, 2015.
- BREM, A.; VOIGT, K. I. **Innovation management in emerging technology ventures** – the concept of an integrated idea management. International Journal of Technology, Policy and Management, v. 7, n. 3, p. 304-321, 2007.
- CARBONE, F.; CONTRERAS, J.; HERNÁNDEZ, J. Z.; GOMEZ-PEREZ, J. M. **Open Innovation in an Enterprise 3.0 framework**: Three case studies. Expert Systems with Applications, v.39, n. 10, p. 8929-8939, 2012.
- CHEN, J. et al. **Visualization of Ideas**: A Valuable Tool for Managing Complex Projects. In: Proceedings of the International Conference on Engineering and Technology, 2022. p. 1-8.
- COOPER, Robert G. **Stage-gate systems**: A new tool for managing new products, Business Horizons, vol. 33, n. 3, p. 44-54, 1990.
- DZIALLAS, Marisa; BLIND, Knut. **Innovation indicators throughout the innovation process**: An extensive literature analysis. Technovation, v. 80, p. 3-29, 2019.
- EBERHARD, Karin. **The effects of visualization on judgment and decision-making**: a systematic literature review. Management Review Quarterly, v. 73, n. 1, p. 167-214, 2023.

FLYNN, M. et al. **Idea management for organizational innovation**. International Journal of Innovation Management, v. 7, n. 5, p. 417-442, 2003.

GABRIEL, A. MONTICOLO; D., CAMARGO, M.; BOURGAULT, M. **Ontology to Represent the Knowledge Domain of a Creative Workshop**. In: 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS). IEEE, p. 618-623, 2016.

GAMA, Fábio; FRISHAMMAR, Johan; PARIDA, Vinit. **Idea generation and open innovation in SMEs: When does market-based collaboration pay off most?**. Creativity and Innovation Management, v. 28, n. 1, p. 113-123, 2019.

GAYNOR, G.H. **Innovator: what does it take to be one?** IEEE Antennas and Propagation Magazine, vol. 43, nº. 3, p. 126–130, 2001.

GIBSON, R.; SKARZYNSKI, P. **Inovação: prioridade nº 1: o caminho para transformação nas organizações**. Rio de Janeiro: Elsevier, 2008.

GIL, A. C. **Como elaborar projetos de pesquisa**. São Paulo: Atlas, 2008.

HANSEN, A. **Innovation: A Continuous Process of Creating and Adopting New Ideas, Products, and Services**. In: Proceedings of the International Conference on Business and Management, 2022. p. 1-10.

HEVNER, A. R. **A three cycle view of design science research**. Scandinavian Journal of Information Systems, v. 19, n. 2, p. 87-92, 2007.

HILBERT, Martin. **Visualização de dados: definição, exemplos e boas práticas**. São Paulo: Novatec, 2016.

HOORNAERT, S.; BALLINGS, M.; MALTHOUSE, E. C.; VAN DEN POEL, D. **Identifying new product ideas: waiting for the wisdom of the crowd or screening ideas in real time**. Journal of Product Innovation Management, v. 34, n. 5, p. 580-597, 2017.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. **Data Clustering: a review**. ACM Computing Surveys, v. 31 nº 3, p. 264-323, 1999.

JAIN, Anil K.; DUBES, Richard C. **Algorithms for Clustering Data**. Englewood Cliffs: Prentice Hall, 1988.

JOLLIFFE, I. T. **Principal component analysis**, Second Edition. Springer, 2002.

KHAN, M. A.; SIDDIQUI, M. S. **A Review of Clustering Methods for Big Data**. ACM Computing Surveys (CSUR), v. 56, n. 6, p. 1-37, 2023.

KOEN, Peter A. et al. **Fuzzy front end: Effective methods, tools, and techniques**. New York, NY: Wiley, 2002.

KONCHADY, Manu. **Text mining application programming**. Massachusetts: Charles River Media, 2006.

KOVACOVA, Maria; MACHOVA, Veronika; BENNETT, Daniel. **Immersive extended reality technologies, data visualization tools, and customer behavior analytics in the metaverse commerce**. Journal of Self-Governance and Management Economics, v. 10, n. 2, p. 7-21, 2022.

LEE, H.; CHOI, K.; YOO, D.; SUH, Y.; LEE, S.; HE, G. **Recommending valuable ideas in an open innovation community**: A text mining approach to information overload problem. Industrial Management & Data Systems, v. 118, n. 4, p. 683-699, 2018.

MARTINEZ-TORRES, Rocio; OLMEDILLA, Maria. **Identification of innovation solvers in open innovation communities using swarm intelligence**. Technological Forecasting and Social Change, v. 109, p. 15-24, 2016.

MAHROUM, Sami; AL-SALEH, Yasser. **Towards a functional framework for measuring national innovation efficacy**. Technovation, vol. 33, n° 10–11, p. 320-332, 2013.

MUNZNER, T. **Visualization analysis and design**. CRC Press, 2014.

NORRIS, S. **Data Visualization**: An Effective Technique for Managing Large Datasets and Identifying Insights for Idea Management. In: Proceedings of the International Conference on Data Science and Big Data Analytics, 2022. p. 1-8.

PADILLA, Lace M. et al. **Decision making with visualizations**: a cognitive framework across disciplines. Cognitive research: principles and implications, v. 3, n. 1, p. 1-25, 2018.

PEFFERS, K.; TUUNANEN, T.; ROOSTHUIZEN, M.; CHATTERJEE, S.; SIMPSON, M. **A design science research methodology for information systems research**. Journal of Management Information Systems, v. 24, n. 3, p. 45-77, 2007.

POLHEMUS, Ashley et al. **Data visualization for chronic neurological and mental health condition self-management**: systematic review of user perspectives. JMIR mental health, v. 9, n. 4, p. e25249, 2022.

POVEDA, G.; WESTERSKI, A.; IGLESIAS, C. A. **Application of semantic search in Idea Management Systems**. International Conference for Internet Technology And Secured Transactions, 2012, vol., no., p.230 - 236, 10-12 Dec. 2012.

RAMOS, J. **Using TF-IDF to determine word relevance in document queries**. In: Proc. of the first int. conf. on machine learning, 2003.

SALDIVAR, Jorge et al. **Idea Management Communities in the Wild**: An Exploratory Study of 166 Online Communities. In: 2016 International Conference on Collaboration Technologies and Systems (CTS). IEEE, 2016. p. 81-89.

SCHUMPETER, J.A. **The theory of economic development**: an inquiry into profits, capital, credit, interest and the business cycle; translated from the German by Redvers Opie. Cambridge, MA: Harvard University Press, 1949.

TARTUCE, T. J. A. **Métodos de pesquisa**. Fortaleza: UNICE–Ensino Superior, 2006.
VANDENBOSCH, B.; SAATCIOGLU, A.; FAY, S. Idea management: a systemic view. *Journal of Management Studies*, Malden, v. 43, n. 2, p. 259-288, 2006.

WANG, Deqiang; GUO, Danhuai; ZHANG, Hui. **Spatial temporal data visualization in emergency management**: a view from data-driven decision. In: Proceedings of the 3rd ACM SIGSPATIAL Workshop on Emergency Management using. 2017. p. 1-7.

WESTERSKI, Adam; IGLESIAS, Carlos Angel. **Exploiting Structured Linked Data in Enterprise Knowledge Management Systems**: An Idea Management Case Study. EDOCW. p. 395-403, IEEE Computer Society, 2011.

WONG, Dona M.; KATZ, Irving. **The Wall Street Journal Guide to Information Graphics**: The Dos and Don'ts of Presenting Data, Facts, and Figures. New York: W. W. Norton & Company, 2010.

ZHANG, P.; LI, Y.; ZHANG, D.; YANG, X. **A Survey on Clustering Algorithms for Big Data**. *IEEE Access*, v. 11, p. 48833-48858, 2023.